# The Role of the Market and Competition in M.B.A. Degree Program Rankings: The Reliability and Validity of Rankings of M.B.A. Programs

A Dissertation
SUBMITTED TO THE FACULTY OF THE
UNIVERSITY OF MINNESOTA
BY

Olena Glushko Horner

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Dr. Darwin D. Hendel, Adviser

June 2016

# ACKNOWLEDGMENTS

I would like to thank the many people who have assisted me in this study with their time, knowledge, and resources. First and foremost, I owe a debt of gratitude for the guidance, support, wisdom, and optimism which my advisor, Professor Darwin Hendel, has offered me. I could not have imagined a better advisor and mentor for my Ph.D. studies. A brilliant scholar and great teacher, whose expertise goes beyond the narrow field of higher education, Professor Hendel provided me with a wealth of help, ranging from advice on complex statistical procedures to intricacies of writing about the limitations of your research. Professor Hendel is the most caring advisor who always helped explore and find job opportunities, wrote letters of recommendation, and kindly reminded me of the requirements for degree progress if I was distracted from my Ph.D. studies by other commitments in my life. Darwin is the advisor who knows your son's name, hobbies, and talents and understands the challenges of parenting and going to graduate school at the same time.

I would also like to acknowledge the dedication and support of my committee members, Professor Darwin Hendel, Professor David Weerts, Professor Robert delMas, Dr. Jeremy Hernandez, and Professor Melissa Anderson, and thank them for sharing their expertise with me along the way. I want to say special thanks to Professor delMas for being a wonderful and caring statistics teacher who did not scare his students out of pursuing quantitative research. Professor delMas provided me with valuable insights into statistical procedures and feedback without which this dissertation would not have been completed.

I want to express my gratitude to the raters who participated in this dissertation research for their contribution, support, and insightful comments and suggestions.

I am grateful to the Office of Institutional Research that provided a supportive environment for my education and offered plenty of opportunities to grow as a researcher. I want to say thank you to Dr. Daniel Jones-White, Dr. Ronald Huesman, Jr., and John Kellogg for being great mentors.

My sincere thanks also goes to the advisors of my first Ph.D. dissertation in Ukraine, Professor Natalya Tigranovna Pakhsarayn and Professor Tatyana Nikolaevna Potnitseva. Although my first dissertation was in English Language and Literature, Natalya Tigranovna and Tatyana Nikolaevna taught me the foundations of research skills that were transferable to other fields.

There are no words to express the enormity of gratitude I feel for my family for their support, inspiration, and patience. From the bottom of my heart, I want to thank my husband, Patrick, for his enormous love, inexhaustible patience, overwhelming support, and help in finding research literature and factual and statistical information for my dissertation. I am grateful for the love of my son, Max, and my niece, Veronika, who are the rays of sunshine in my life. I am especially indebted to my parents, Nina and Alexandr, for their immeasurable love, support, and sacrifices for my education and success. I am thankful to my sister, Tatyana, who loves me dearly and supports in life. I am in deep gratitude to my grandparents, Petr and Maria and Ivan and Vera, who loved me dearly, supported me in all my endeavors, and raised me a hard-working person. I am very thankful to my extended family here, especially to my mom-in-law, Joanne, my sister-in-law Laura, and my brother-in-law Patrick, who made me feel at home in this

country. I also want to thank my best friend and former colleague Tatyana Lyaskovets for her support, professional advice, and the bonds of friendship which even the Atlantic Ocean between us cannot break.

## DEDICATION

I dedicate this dissertation

to the love of my life, my husband, Patrick;

to the sunshine of my life, my son, Max, and my niece, Veronika;

to my loving parents, Nina and Alexandr;

to my beautiful sister, Tatyana;

to my wonderful grandparents, Petr and Maria.

## Abstract

In light of increasing popularity of rankings of M.B.A. programs among prospective students, employers, and the general public, this dissertation research seeks to explore five ranking systems of full-time M.B.A. programs in the United States: *U.S. News & World Report, Bloomberg Businessweek, The Financial Times, The Economist,* and *Forbes*, from four different perspectives. First, the study assesses the quality of the five full-time M.B.A. program rankings through the lens of the Berlin Principles on Ranking of Higher Education Institutions. Further, the study compares the resulting ranking of M.B.A. rankings, which is based on the objective quality assessment of these ranking systems, with students' evaluation of the importance of various M.B.A. rankings for their decision to enroll as reported in the 2012 mba.com Prospective Students Survey (Graduate Management Admission Council, 2012b). The differences in ranking of M.B.A. rankings produced by prospective students and higher education professionals bring into focus the necessity for a better alignment of priorities and values of the two groups.

Second, by exploring first-order autocorrelation of the error terms for the four rankings of M.B.A. programs in the United States: *U.S. News & World Report, Bloomberg Businessweek, The Financial Times,* and *Forbes*, this study demonstrates that methodological changes made to these rankings between 2000 and 2013 have enhanced their reliability in terms of predictability, which suggests that ranking entities are responsive to criticism and have made an effort to improve their methodologies. The results of the correlational analysis of the five full-time M.B.A. ranking systems in the

United States across 14 years strongly support the hypothesis about the stability of ranking scores, which lends credence to the reliability of the rankings.

The examination of the degree of convergent and discriminant validity of the *Bloomberg Businessweek* full-time M.B.A. program rankings using multitrait-multioccasion confirmatory factor analysis allowed developing evidence for the validity of measures that are used in the rankings of M.B.A. programs to determine the hypothetical construct of quality of an M.B.A. program.

Third, by examining the anchoring effect of rankings and reputation in higher education, this study demonstrates that the current university reputation has a strong impact on the future reputation of its business school. These associations solidified from year to year which suggests practical implications for collaboration between university and business school administrators.

Finally, the specific attribute of this study consists in exploring how full-time M.B.A. program rankings affect funding streams to higher education institutions. Disentangling the differential impact of the rank of a business school on R&D funding coming to the university demonstrates positive associations between business school rankings and funding coming from federal government and nonprofit organizations and negative associations between funding coming from state and local government and institutional funds.

*Keywords:* reliability, predictability, first-order autocorrelation, stability, construct validity, convergent and discriminant validity, rankings of M.B.A. programs, anchoring effect, reputation, funding streams to higher education institutions

Table of Contents

List of Tables

List of Figures

# CHAPTER I: INTRODUCTION

In this chapter, I delineate the problems that business education faces and provide the rationale for public accountability of business schools in public and private universities. I concentrate on the emerging role of rankings as a public accountability tool for business schools and their impact on prospective students and corporate recruiters. In light of increasing popularity of rankings of M.B.A. programs among prospective students, employers, and the general public, I emphasize the need to assess the reliability and validity of the rankings of M.B.A. programs and examine their impact on higher education institutions in the United States.

## The Need for Public Accountability of Business Schools

Although business school enrollments have soared and business schools have introduced several innovations to fill the needs of a changing world, a growing number of studies cast doubt on the value of graduate business education (Bailey & Ford, 1996; Bennis & O'Toole, 2005; Mintzberg, 1996, 2004; Pfeffer & Fong, 2002, 2003, 2004; Podolny, 2009), which raises an issue of assessing institutional quality. Researchers (Bennis & O'Toole, 2005; Mintzberg, 1996, 2004; Pfeffer & Fong, 2002, 2004; Podolny, 2009) criticized business schools for inadequate curriculum, which forces business companies to "offer extensive training and development to new hires" (Pfeffer & Fong, 2004, p. 1508), ineffective teaching methods, and the quality of faculty who lack practical experience in the corporate world. Bennis and O'Toole (2005), for example, noted that "today it is possible to find professors of management who have never set a foot inside a real business, except as customers" (p. 101). Researchers (Bailey & Ford,

1996; Bennis & O'Toole, 2005; Mintzberg, 1996, 2004; Pfeffer & Fong, 2002, 2004) also complained about the relevance of business school research. As Bailey and Ford (1996) explained it, "business schools appeal to one another as scholarly communities through a plethora of academic journals that are utterly divorced from the challenges of everyday management" (p. 8). Finally, researchers (Bennis & O'Toole, 2005; Podolny, 2009) voiced concern about failure of M.B.A. programs to inculcate students with norms of ethical behavior. Because of these negative tendencies and the fact that "the [MBA] industry is overbuilt" (Hamada as cited in Gaddis, 2000, para. 2), researchers suggest requiring "systematic assessments of business school products and more attention to the competitive environment" (Pfeffer & Fong, 2002, p. 93).

Consequently, the high tuition costs, on the one hand, and the overcrowded industry of M.B.A. education, on the other hand, which does not always meet students' expectations of high quality and value added, stipulate the need for public accountability. In this respect, media rankings of M.B.A. programs, which both provide information about a business school's performance and allow analyzing it from a comparative perspective, seem to meet the demands for public accountability. Moreover, rankings bring both the public and private sectors of business education into the limelight. Unlike public higher education institutions that are more subject to public scrutiny through various accountability methods, such as performance-based funding, state governance, federal and state policies, etc. (Zumeta, 2001), private higher education institutions face less pressure from external constituents. Because private business schools charge higher tuition prices, they should account to the general public for the high cost of their M.B.A.

programs. Therefore, as a result of tremendous expansion, business education and especially "MBA programs have attracted considerable societal interest over the last couple of decades and have become, rightly or wrongly, the focus of numerous rating systems and huge investment" (Friga, Bettis, & Sullivan, 2003, p. 234).

Besides the problem of the inadequate quality of M.B.A. education and high tuition costs, the student debt crisis and the problem of return on investment in an M.B.A. degree have also become important reasons for greater accountability of business schools. Although lucrative earnings of M.B.A. graduates are well publicized and ranged, for example, from $142,501 in the average starting salary and bonus for the Harvard Business School graduates to $41,400 for the Graduate School of Business at Belmont University graduates in 2012 (*U.S. News & World Report*, 2013a), business schools are reluctant to disclose statistics about their student debt on their websites and in their promotional materials. According to data reported by 94 business schools in an annual *U.S. News & World Report* survey in 2011, on average, full-time M.B.A. students who borrowed for business school finished with $47,743 in debt (Lytle, 2012). On the high end, on average, M.B.A. student debt reached $96,805 (the Fuqua School of Business at Duke University) in 2011 (Lytle, 2012) and increased to $105,782 (the Leonard N. Stern School of Business at New York University) in 2012 (*U.S. News & World Report,* 2013b). On the low end, M.B.A. programs at some schools left students owing as little as $8,000, on average, in 2011 (Hopkins, 2012). While it is tempting to speculate that high-ranked private business schools leave M.B.A. graduates with the most student debt, M.B.A. students at public business schools have become increasingly reliant on student

loans to finance their education. In 2012, M.B.A. students at the University of North Carolina-Chapel Hill, for example, graduated with $82,784 in debt, on average, while M.B.A. recipients at the University of Minnesota-Twin Cities left with an average student loan burden of $61,916 (*U.S. News & World Report,* 2013b).

High costs of M.B.A. education, which lead to heavy indebtedness, aggravated by inability to find employment and low return on investment, raise the question of the quality of business schools and their marketing strategies that may mislead prospective students. According to the data provided by *U.S. News & World Report* (2013a, 2013b), the Jack C. Massey Graduate School of Business at Belmont University, for example, charged $46,050 for its full-time M.B.A. program, and its students graduated with $25,500 in debt, on average, in 2012. At the same time, the average starting salary and bonus of its graduates were only $41,400. Moreover, at graduation, only 26.1% of graduates of the full-time M.B.A. program at Belmont University were employed. This example demonstrates the importance of the quality of information that helps prospective M.B.A. students in their decision-making while they navigate through the promotional materials of 1,120 institutions that award M.B.A. degrees in the United States (U.S. Department of Education, National Center for Education Statistics, 2011). In this respect, rankings of M.B.A. programs may serve as a valuable search tool for prospective students that does not only provide information about various aspects of business schools' performance, but also allow prospective students to compare M.B.A. programs based on certain criteria tailored to their individual goals and available financial resources.

The idea of rankings serving as a public accountability tool echoes that of a recent government initiative "to create the nation's first federal college ratings system to help students make smarter choices about higher education" (Liberto, 2013, para 1). The proposed ratings system planned to rate colleges "based on opportunities -- are they helping students from all backgrounds to succeed -- and on their outcomes" (Obama as cited in Liberto, 2013, para 4), such as graduation rates, debt rates, and access for low-income students. Although many groups of stakeholders in higher education opposed the ratings plan, and the federal government eventually abandoned the idea of creating a college ratings system, it still plans to develop "its ratings replacement – an online college comparison tool" (Stratford, 2015, para 3). Unlike the existing federal government's tools, such as College Navigator and the College Scorecard, which show "a wealth of information" (Cook as cited in Stratford, 2015, para 30), but are not user-friendly, the new consumer tool will provide students and their families with "comparative information on outcomes and college value" (Duncan as cited in Stratford, 2015, para 13).

**The Impact of M.B.A. Rankings on Prospective Students and Corporate Recruiters**

According to the 2012 mba.com Prospective Student Survey administered by the Graduate Management Admission Council, published rankings were the sixth most frequently consulted resource among the top 20 resources that prospective students rely on when making their application decisions (p. 26). Two out of five prospective students (41%) reported that they consulted the published school rankings, and 31% of these respondents considered the published rankings "extremely influential" (Graduate

Management Admission Council, 2012b, p. 27). When asked to rank which ranking publications they deemed extremely influential in their decision-making compared to a school's website, respondents rated *The Financial Times* (17%), *U.S. News & World Report* (16%), and *Bloomberg BusinessWeek* (16%) only slightly lower than school websites (24%). *The Wall Street Journal* (11%), *The Princeton Review* (8%), *Forbes* (8%), *The Economist* (5%), *Canadian Business* (2%), and *Asia Inc.* (2%) were considered to be much less influential in respondents' decision to apply to a particular program (Graduate Management Admission Council, 2012b, p. 27).

Although prospective students ranked school websites (88%) as the most frequently consulted information source (Graduate Management Admission Council, 2012b, p. 26) and considered them "about 50% more influential than the most highly regarded school ranking publication specified, namely, *The Financial Times*" (Graduate Management Admission Council, 2012b, p. 27), they mentioned school rankings among the top 10 types of information which they seek when exploring a business school's website. Specifically, 68% of respondents indicated that they sought information about school rankings on a school's website, and only three items related to respondents' school selection criteria, such as application requirements (73%), cost of attending (71%), and admission criteria (70%), were sought more frequently on a school's website (Graduate Management Admission Council, 2012b, p. 28).

Prospective M.B.A. students' interest in the information about school rankings is not surprising, given the attention that recruiters pay to a business school's reputation and rankings when selecting applicants to interview for a job. A recent study "Eye Tracking

Online Metacognition: Cognitive Complexity and Recruiter Decision Making", conducted by TheLadders (2012), revealed that recruiters spend only six seconds reviewing an applicant's resume. Moreover, the study demonstrated that recruiters spent almost 80% of their resume-review time on only six data points: name, current title/company, previous title/company, previous position start and end dates, current position start and end dates, and education. These findings were congruent with the results of the Corporate Recruiters Survey administered by the Graduate Management Admission Council in 2012. The top five factors selected by the corporate recruiters when they considered choosing applicants to interview included job function at past work experience (61%), industry of past work experience (51%), years of past work experience (44%), graduate degree area of study (40%), and school/institution at which applicants earned their degree (36%) (Graduate Management Admission Council, 2012a, p. 23). At the same time, when corporate recruiters were asked to rate 21 key school criteria[1] that they use when deciding where to conduct on-campus recruitment, school ranking, with an average rating of importance of 3.28, was ranked number two after the quality of the students (4.02) and was followed by the quality of the curriculum (3.16) and reputation of the school (3.10) (Graduate Management Admission Council, 2012a, p. 28).

The findings of the surveys demonstrate the importance of rankings of business schools for both prospective M.B.A. students in their application decisions and corporate recruiters in their recruitment decision-making. The increased reliance of prospective

---

[1] The school criteria were measured on a scale ranging from 1 (*least important*) to 5 (*most important*) (Graduate Management Admission Council, 2012a, p. 28).

students and potential employers on the information provided in rankings raises the question of the quality of different ranking systems that seek to assess the quality of business schools based on different institutional characteristics. As Usher and Savino (2006) framed it, "an institution's ranking is largely a function of what the ranking body chooses to measure" (p. 31). Given the fact that "quality in education is a highly contested notion" (Usher & Savino, 2006, p. 9), the need for assessing the reliability and validity of different rankings of M.B.A. programs in the United States has become the rationale for this research.

**From Theory to Practice: A Framework for the Elaboration and Dissemination of Rankings**

There has been a gradual shift over the previous decade in public, academic, and business school administrators'[2] opinion away from strong criticism and calls to boycott

[2] In the latest issue of *Poets & Quadrants*, Everitt (2014) quoted business school deans on the conundrums of rankings. Although business school deans continued to criticize the rankings, the need to learn from rankings formed the leitmotif of the stories. Idalene Kesner, dean of Indiana University's business school, for example, admitted that "We have woken up to the fact that we need to focus on the rankings in and of themselves. Rankings are a critical part of student decisions as to what school they will go to, and rankings allow a school to reflect on what they do well and areas where they can improve" (as cited in Everitt, 2014, p. 15). The opinions of deans of European business schools echoed those of their American colleagues. Erik Schlie, associate dean of M.B.A. programs at IE Business School in Spain, claimed that "Rankings are

the rankings to their acceptance as "a market-based perspective that can complement the work of government, accrediting authorities, and independent review agencies" (The International Ranking Expert Group (IREG), 2006, p. 2). Recent research literature (Glick, 2008; Hou, Morse, & Shao, 2012; Stolz, Hendel, & Horn, 2010), conference presentations (Rauhvargers, 2013, April 11-12), debates (EUA Annual Conference, 2013, April 11-12), and even media corporations (Thomson Reuters, 2010) focused on how ranking systems could "become more sophisticated benchmarking exercises" (Haworth, 2013, para. 1).

On a practical level, responding to the need for a set of best practices to guide the design of new rankings and to assess the existing rankings, which was articulated in the academic literature (Altbach, 2006; Clarke, 2002; Merisotis & Sadlack, 2005; Proulx, 2006), the UNESCO European Center for Higher Education (UNESCO-CEPES) in Bucharest and the Institute for Higher Education Policy in Washington, DC founded the International Ranking Expert Group (IREG) in 2004. In 2006, IREG came up with the Berlin Principles on Ranking of Higher Education Institutions (IREG, 2006) that set a framework for the assessment of national, regional, and global rankings aiming at creating "a system of continuous improvement and refinement of the methodologies used to conduct these rankings" (IREG, 2006, p. 1). In 2006, the IREG Observatory on Academic Ranking and Excellence ("IREG Observatory") was created on the basis of

---

undoubtedly an important signaling tool, and hence, they fuel the brand of the school with prestige. It's like having a respected, unbiased third opinion about the relevant players in the market" (as cited in Everitt, 2014, p. 17).

IREG in order to assess the quality of existing ranking systems of higher education institutions worldwide. In 2011, "IREG Observatory" undertook the IREG Ranking Audit initiative, the ultimate goal of which was "to improve the overall quality of rankings" (IREG, 2011, p. 5). The "IREG Observatory" methodology for assessing ranking systems, which is specified in the IREG Ranking Audit Manual (IREG, 2011), is based on the Berlin Principles on Ranking of Higher Education Institutions. On a voluntary basis, rankings can undergo an audit process[3], and those ones that pass IREG's evaluation will be "entitled to use the quality label and corresponding logo *"IREG approved"* (IREG, 2011, p. 5). "IREG Observatory" hopes that negative audit decisions will provide an incentive for ranking organizations to improve the quality of their rankings. As a result, conceived as "a public-responsibility initiative" (IREG, 2011, p. 12), the IREG Ranking Audit initiative has become an important endeavor on the part of the international community of higher education professionals, researchers, and policy-makers that provides a framework for a multifaceted assessment of ranking systems.

**Research Design**

The framework for this research was based on the idea of a comprehensive and

---

[3] To date, one international university ranking system (*QS World University Rankings*) and two national university ranking systems (*Perspektywy University Ranking* (Poland) and *Center for Higher Education University Ranking* (Germany)) have passed the IREG Ranking Audit and received IREG Approved certificate (IREG, 2015a). IREG also publishes the "IREG Inventory of National Rankings" on its website and highlights rankings with the "IREG Approved" label (IREG, 2015b).

multifaceted audit of ranking systems suggested by the IREG Ranking Audit initiative. Accordingly, this study seeks to explore five ranking systems of full-time M.B.A. programs in the United States: *U.S. News & World Report, Bloomberg Businessweek, The Financial Times, The Economist,* and *Forbes*, from four different perspectives. First, the study intends to assess the quality of the five full-time M.B.A. program rankings through the lens of the Berlin Principles on Ranking of Higher Education Institutions. Second, the study aims to make a comprehensive assessment of the reliability and validity of the five full-time M.B.A. program rankings mentioned above. Third, this study seeks to examine the relationship between a full-time M.B.A. program ranking and an overall ranking of an institution. Finally, this study explores how full-time M.B.A. program rankings affect funding streams to higher education institutions.

**Research Questions**

The following research questions guided this study:

1. To what extent do objective assessments of the quality of the five full-time M.B.A. rankings in the United States, as determined by the Berlin Principles on Ranking of Higher Education Institutions, correlate with how the prospective students evaluate the M.B.A. ranking systems?
2. What is the reliability and construct validity of the five full-time M.B.A. rankings in the United States?
3. Is the measure of quality of full-time M.B.A. programs independent of the measure of quality of higher education institutions?

4. How do rankings of full-time M.B.A. programs affect funding streams to institutions of higher education?

# CHAPTER II: REVIEW OF THE LITERATURE

This chapter is devoted to the review of the literature on rankings of M.B.A. programs in the United States. The literature review addresses the role of rankings of M.B.A. programs as a means of public accountability, summarizes empirical findings on the influence of rankings on business schools' policies and operations, provides a brief description of the five major media rankings of full-time M.B.A. programs in the United States, and delineates contemporary problems in rankings methodologies. The chapter concludes with the outline of the research gap in the study of M.B.A. rankings in the U.S.A., the exploration of which accounts for the contribution of the present study to existing research.

## The Institutional Effects of Rankings of Business Schools

**Positive effects of ranking systems.** Rankings of business schools mirror the same "love-hate relationship" (Bradshaw, 2007, p. 54) that is characteristic of the relationship between rankings and higher education institutions overall. This ambivalent relationship is grounded in the competitive environment, which "in great part was created by the rankings" (Corley & Gioia, 2000, p. 321). Nevertheless, the unique characteristic of this competition among business schools, as Corley and Gioia (2000) argued, is "the strong belief (and acceptance) by the players that the rankings game is a contrived, socially constructed, and socially sustained competition" (p. 321). "The social construction of a competition" (Corley & Gioia, 2000, p. 320) accounts for a number of advantages of the rankings for business schools. According to Gioia and Corley (2002), the first advantage of rankings lies in the fact that "competition improves the breed" (p.

109). Gioia and Corley (2002) argued that competition had improved M.B.A. programs and had made them more relevant to the modern business world.

The acknowledgement of the positive impact of competition, which M.B.A. rankings provide for business schools, is echoed in research on the impact of rankings in other professional fields. In law education, for example, Korobkin (1998) argued that "Rankings create status; status creates competition for status; and competition creates activity that can be put to productive use" (p. 417). Consequently, in Korobkin's (1998) opinion, if rankings of law schools intensify competition and stimulate the production of legal scholarship, they can "encourage the creation of public goods that otherwise would be inefficiently underproduced by the law school marketplace" (p. 404).

The second advantage of rankings, in Gioia and Corley's (2002) view, consists in the fact that business schools face real competition and learn how to survive in a competitive environment, which enriches business programs. As Gioia and Corley (2002) noted with a touch of irony in their voice, business schools were "teaching *about* competition without *knowing* competition firsthand" a decade ago (p. 109). At present, a "high-velocity environment", which is characterized by "rapid and discontinuous change in demand, competitors, technology, or regulation, so that information is often inaccurate, unavailable, or obsolete" (Eisenhardt & Bourgeois, 1988, p. 738), forces business schools to change rapidly and gives them the opportunity to test the market theories and strategies that they teach. Within this context, business schools have started to pay closer attention to rankings results in order to identify students' expectations of an M.B.A. program. Samy, Hodgson, and Moulton (2007), for example, reviewed the criteria of *The*

*Economist* rankings of full-time M.B.A. programs in 2006, as well as the program structure of the top eight business schools in the UK, and proposed a model for an M.B.A. program which is designed "based on the criteria that the EIU[4] value in their rankings" (p. 92). Samy et al. (2007) concluded that "The recommended framework for an MBA programme takes into account the criteria as is applied by the EIU in their ranking of MBA schools globally" (p. 79).

The third advantage of rankings, with which most scholars are in complete agreement (Bickerstaffe & Ridgers, 2007; Gioia & Corley, 2002), is that they have made business schools more transparent and accountable to the public. The surge of interest in accountability in the public higher education sector engendered various accountability initiatives at the state level that impact higher education institutions through performance funding (John, Kline, & Asker, 2001; Zumeta, 2001). Burke (as cited in Zumeta, 2001) conducted extensive research on various models of performance funding and provided a taxonomy of accountability measures that includes input, process, output, and outcome measures. In light of Burke's (as cited in Zumeta, 2001) study, it is possible to conclude that most rankings criteria in the five most influential media rankings of M.B.A. programs in the U.S.A. (*Bloomberg Businessweek, The Financial Times, Forbes, The Economist,* and *U.S. News & World Report*) coincide with the input and outcome accountability measures identified by Burke (as cited in Zumeta, 2001, pp. 169-170).

---

[4] EIU stands for *Economist* Intelligence Unit which is an independent business within The Economist Group.

Burke's list of the most common performance indicators, which are used in states that employ performance-based funding (as cited in Zumeta, 2001, p. 170), includes "undergraduate retention and graduation rates", "professional licensure test scores or pass rates", "transfers from community colleges to the baccalaureate programs", "use of technology or distance learning in teaching", "faculty teaching load measures", "credits at graduation or time required to obtain a degree", "faculty/staff diversity indicators", "indicators of a typical student's (economically) feasible choices among colleges", "job placement after graduation", preparation levels of entering students", "noninstructional cost as a share of all costs", "program duplication across campuses", survey results on satisfaction (of alumni or employers)", "sponsored research funds obtained", "student learning measures (test scores)", and "workforce training/development indicators".

Ranking systems of M.B.A. programs, with the exception of the reputational component, demonstrate a considerable overlap with accountability measures. Rankings criteria in the context of business schools that include student and faculty characteristics, such as undergraduate GPA, student GMAT scores, acceptance rate to program, international faculty and students, women faculty and students, and faculty with doctorate, correspond to the accountability input measures. The outcome measures that provide an indication of "the social value of the output" (Zumeta, 2001, p. 169) include surveys of students and recruiters, employment rates, starting salaries, and growth in salary after receiving an M.B.A. As a result, the media's choice of rankings criteria, which scholars (Bickerstaffe & Ridgers, 2007; Corley & Gioia, 2000; DeAngelo, DeAngelo, & Zimmerman, 2005; Dichev, 1999; Fisher, Kiang, & Fisher, 2007; Rindova,

Williamson, Petkova, & Sever, 2005; Tracy & Waldfogel, 1997) claim in unison to be a poor proxy for assessing the quality of business schools, do not seem to be completely arbitrary. As Gioia and Corley (2002) noted, rankings are "not all just artifice" (p. 110) and "are not absurd" (p. 110). The media have borrowed these indicators of quality from the public sector of higher education, applied them to rank business schools from both public and private sectors nationwide, and made them readily available to the public.

Accountability and transparency bring about the fourth advantage of rankings in the context of business schools that "introduced a kind of democracy to the management development industry" (Bickerstaffe & Ridgers, 2007, p. 65). The M.B.A. sector is "a crowded field" (Lorange, 2005, p. 784), and less prominent business schools do not have enough resources to gain publicity in this competitive landscape. In this respect, the positive impact of rankings on business schools consists in the fact that rankings "have provided a much higher profile for many smaller or less well-known business schools, both within their own countries and, more importantly, globally" (Bickerstaffe & Ridgers, 2007, p. 65).

The fifth advantage of rankings considers accountability and transparency from a different perspective. In his study of the impact of reputational rankings on organizational change, Martins (2005) asserted that rankings provide business schools' managers with "feedback about their organization's relative standing in its industry" (p. 702). Martins (2005) concluded that business schools with a "perceived identity-reputation discrepancy" (p. 703) are more prone to undertake organizational change in response to rankings. As a result, transparency in the field has become "a central, driving

force behind the evolution of competitive success and failure" (Barnett & Hansen, 1996, p. 139), which is known as "the Red Queen" effect (Barnett & Hansen, 1996, p. 139) in evolutionary theory. "The Red Queen" effect posits that an organization that faces competition searches for new strategies to improve performance. Successful strategies increase the organization's competitiveness and at the same time trigger "learning in its rivals - consequently making them stronger competitors and so again triggering learning in the first organization" (Barnett & Hansen, 1996, p. 139). As a result, transparency of the field has improved the quality of business schools and made them "more strategic in deciding or designing the domains" (Gioia & Corley, 2002, p. 110) in which they want to dominate because "it is hard to maintain a competitive advantage when you develop a practice that is there for all to see" (Gioia & Corley, 2002, p. 110).

The final positive attribute of rankings, as Gioia and Corley (2002) argued, lies in the fact that rankings benefit prospective students and recruiters "who previously had to cope with a great deal of managed translucency on the part of business schools" (p. 110). At present, rankings serve as an approximate guide in assessing business schools' quality based on certain criteria.

This argument about the positive attribute of rankings of business schools for prospective students and employers resonates with scholars' opinions from other professional fields. In law education, for example, Berger (2001) posited that rankings provide prospective students with "quick" and "convenient" information and allow comparing law schools "with respect of important and relevant criteria" (p. 497). Moreover, Korobkin (1998) asserted that law school rankings "serve a coordination

function" (p. 408) that implies matching the best students with the best institutions, and, subsequently, with the best employers.

Overall, Corley and Gioia (2000) concluded that "the existence of the rankings actually provides a strong impetus for improving the quality of business schools" (p. 330). Organizational theorists argue that "legitimacy is an intangible asset" (Rao, 1994, p. 29) that impacts the ability of organizations to acquire tangible resources and influences their survival (Dowling & Pfeffer, 1975; Pfeffer & Salancik, 1978). Various empirical studies demonstrated how "certification contests legitimate organizations, generate status ordering, and create favorable reputations" (Rao, 1994, p. 29) in the corporate world. Rao (1994), for example, showed that victories in certification contests improved organizations' survival in the automobile industry. Fombrun (1996) demonstrated that higher-ranked organizations were able to charge higher prices for their products. Wilson (1985) revealed that certification by established auditors provided firms with better access to capital. Hayward and Boeker (1998) asserted that reputation led to greater professional autonomy in investment banking. Finally, a number of researchers (Fombrun & Shanley, 1990; Podolny, 1993; Roberts & Dowling, 2002; Smith, Smith, & Wang, 2010) reported a positive relationship between a company's reputation and its financial performance. In a related vein, in the higher education sector, Wiley and Zald (1968) demonstrated that accreditation performed a function of legitimization of organizations, which is similar to certification contests and ratings in the corporate world, and helped educational institutions create image, acquire resources, and increase their chances of survival. Therefore, at present, if "properly formulated and

used" (Corley & Gioia, 2000, p. 330), rankings can have a positive impact on business schools.

**Negative effects of ranking systems.** The studies that criticize rankings of business schools are more numerous than those that try to find some positive attributes of the ranking systems. Some researchers have adopted a very negative stance on ranking systems and asserted that "the rankings have damaged the business school industry" (Policano, 2001, p. 40) and are "fundamentally flawed" (Morgeson & Nahrgang, 2008, p. 36). In a similar vein, DeAngelo, DeAngelo, and Zimmerman (2005) claimed that "U.S. business schools are locked in a dysfunctional competition for media rankings that diverts resources from long-term knowledge creation, which earned them global pre-eminence, into short-term strategies aimed at improving their rankings" (p. i). DeAngelo, DeAngelo, and Zimmerman's (2005) criticism of business school rankings goes to extremes in depicting the dangers and deficiencies of the rankings. Nevertheless, the overall negative discourse on business school rankings provides constructive criticism, which both business schools and media rankings can use to improve their quality.

The first disadvantage of rankings, about which the majority of scholars agree, is the Matthew effect (Merton, 1968) that implies that higher-status organizations receive greater rewards than lower-status organizations for fulfilling an identical task (Podolny, 1993). In the higher education sector, the Matthew effect leads to a situation where "marginal differences in performance lead to large differences in reputation and resources" (Bastedo & Bowman, 2010, p. 180). Similar to the effects of rankings on the ability of higher education institutions to garner resources, moving up in the rankings

allows business schools to charge a premium price, attract high quality applicants, obtain resources from alumni and corporate donors, and increase the number of prestigious companies that recruit their graduates (Argenti, 2000; Corley & Gioia, 2000; Eberhardt & Moser, 1997; Rindova et al., 2005; Safón, 2007). Cheslock and Gianneschi (2008) found that an institution's business school ranking might also be a strong predictor of giving revenues for the entire institution. As a result, as Corley and Gioia (2000) concluded, "the snowball effect" of the business school rankings "promotes a 'rich-get-richer and poor-get-poorer' cycle and creates a 'Catch 22' trap from which it is difficult to extricate oneself, win or lose" (p. 329).

Empirical research in the field of legal education (Sauder & Lancaster, 2006) demonstrated the same effects of rankings on external audiences and consequences for law schools. As Sauder and Lancaster (2006) found, law schools with higher ranks receive a higher number of applicants: an increase in a rank in the first tier leads to an increase in the number of top applicants and the percentage of students who matriculate. At the same time, an increase in one rank leads to a reduction in the percentage of students who are accepted. Sauder and Lancaster (2006) concluded that "the existence of such a spiral, negative or positive" (p. 117), produced by the rankings, can "*create* new forms of inequality rather than simply reflect preexisting inequalities among schools" (p. 117). Thus, the similarities in research findings about the negative effects of the rankings on business and law schools validate each other and support generalizing their conclusions to other professional fields.

The second negative consequence of competition for media rankings, in scholars' opinion (DeAngelo et al., 2005; Gioia & Corley, 2002; Podolny, 2009; Policano, 2001), is that "rankings have legitimized most business schools' myopic focus on the short term" (Podolny, 2009, p. 65). The array of adverse effects of the short-term "performance drivers" (Podolny, 2009, p. 65) includes deterioration of the quality of teaching (Gioia & Corley, 2002), distortions in the M.B.A. curriculum as a result of too much emphasis on "trendy topics" (DeAngelo et al., 2005, p. 9) and disregard of fundamentals (Gioia & Corley, 2002), changes in student composition, which is skewed toward scholastically superior applicants, and neglect of the benefits of a diverse student pool for the educational outcomes (DeAngelo et al., 2005; Gioia & Corley, 2002). The prominence of the M.B.A. media rankings has caused "serious misallocations of resources" (Policano, 2001, p. 40) that impact various "core identity dimensions" (Elsbach & Kramer, 1996, p. 444) of some business schools and reduce emphasis on undergraduate education, Ph.D. programs, and research mission (AACBS International, The Association to Advance Collegiate Schools of Business, 2005; DeAngelo et al., 2005; Gioia & Corley, 2002). In order to improve their standing in the rankings, some business schools also channel great efforts towards improving student placement services and intensifying public relations campaigns (Argenti, 2000; Gioia & Corley, 2002). As a result, as Morgeson and Nahrgang (2008) noted, the external pressure of "outside groups that do not share the same values or mission" became so strong that these groups "dictate the agenda of business school education" (p. 37).

Business schools are not the only professional schools whose policies and decision-making are distorted by competition for media rankings. As Espeland and Sauder (2007) reported, law schools also tend to redistribute resources "in order to optimize their rank" (p. 25). Espeland and Sauder (2007) documented the following strategies employed by law schools in reaction to rankings: a dramatic increase in spending on marketing and advertising, a change in admission policies with a greater emphasis on LSAT scores, a shift from need-based to merit-based scholarships, and a greater emphasis on collecting placement statistics by career services at the expense of providing career counseling. As a result, Espeland and Sauder (2007) concluded, "Boosting rankings criteria now threatens to overshadow other goals" (p. 29), which reflects "the pragmatic reality of rankings" (Sauder & Lancaster, 2006, p. 116) not only for business and law schools but also for the entire field of higher education.

The third disadvantage of rankings is also the result of enormous external and internal pressure that the ranking competition imposes on business schools. Gioia and Corley (2002) lamented that "With this level of cynicism and competition bred by the rankings, it comes as no surprise that there is some evidence of schools gaming the system" (p. 113). The gaming practices, for example, ranged from assigning best faculty and allocating more resources to the classes that were surveyed by *Businessweek* to direct counseling of students on how to respond to surveys (Gioia & Corley, 2002) in order to "rank the school highly or risk lowering the value of their own degree" (Argenti, 2000, p. 174). This practice became so popular that *Businessweek* threatened to exclude business

schools that are engaged in counseling their students about responses to the rankings (Gioia & Corley, 2002).

Other professional fields are not immune from the temptation to game the rankings either. Some law schools, for example, as researchers (Espeland & Sauder, 2007; Whitman, 2002) noted, try to increase selectivity by encouraging underqualified students to apply, skew admission policies in favor of applicants with higher LSAT scores, force faculty to take research leaves in spring instead of fall in order to decrease the student-faculty ratio, and hire their unemployed graduates temporarily to improve their placement numbers.

Another significant concern about the validity of rankings is the quality of the data, which are provided by business schools. Among the five major media ranking systems of business schools in the U.S., only *The Financial Times* claims that it independently audits the data, which business schools submit (Bradshaw, 2007). In order to influence the rankings, some schools attempt to manipulate the data by "focusing on 'creatively interpreting' the reporting criteria that produce the rankings" (Corley & Gioia, 2000, p. 326). Corley and Gioia (2000) provided the evidence of such manipulation that includes gaming both input and outcome measures. The examples of falsification of the input measures consist of admitting lower quality applicants into a Master of Science program and transferring them to the M.B.A. program later and putting lower quality applicants and international students, whose future income in their home countries is not comparable to MBAs' income in the U.S., into a 'pre-admission class' that is not reported as the 'entering class' (Corley & Gioia, 2000). The evidence of gaming of the outcome

measures includes providing data about only the average bonus for the graduates who receive a bonus instead of reporting the average bonus for the whole cohort (Corley & Gioia, 2000). As Corley and Gioia (2000) concluded, such unethical behavior has far-reaching consequences because "these practices put business schools at the risk of compromising the integrity of the entire professional education industry" (p. 326).

Law schools, as Espeland and Sauder (2007) claimed, demonstrate the same tendency to skew statistics under rankings pressure. Some law schools were caught reporting inconsistent data to *U.S. News & World Report* and the American Bar Association, which is their accrediting body. Other strategies for gaming the rankings, as Espeland and Sauder (2007) reported, involve reducing the size of incoming classes by admitting fewer students into the full-time program and more students into the part-time program, classifying admitted students with lower LSAT scores as "probationary" instead of full-time students, and counting all graduates as employed although some of them do not practice the legal profession. Such misinterpretations and falsifications do not only mislead the public but also create a backlash against law schools themselves. As Espeland and Sauder (2007) noted, gaming the rankings encourages law school administrators "to be more distrustful of their peers" (p. 32), which undermines the values of the academic community.

The final and the most important problem with the rankings of business schools is associated with their methodology (Corley & Gioia, 2000; Dichev, 1999; Gioia & Corley, 2002; Gladwell, 2011; Martins, 2005; Morgeson & Nahrgang, 2008; Policano, 2007; Rindova et al., 2005; Safón, 2007). Brooks (2005) asserted that in measuring university

quality, "a theoretical definition of quality" should be "consistent with its measurement strategy" (p. 4). Nevertheless, as Corley and Gioia (2000) noted, academia has always had a problem with identifying "'objective, bottom-line' measures of performance or effectiveness" (p. 332), and "subjective perceptions of performance" (p. 332) historically dominated the higher education sector. In this respect, Sweitzer and Volkwein (2009) concluded that there are two major factors that underlie the problems in rankings methodologies. The first factor is "doubts about the validity of the ratings as measures of quality", and the second factor is "conflicting views about what constitutes quality in higher education" (Sweitzer & Volkwein, 2009, p. 814).

**Rankings of business schools and models of quality in higher education.** While reviewing literature on quality in academia, Sweitzer and Volkwein (2009) classified the studies into four groups or "models of quality in higher education" (p. 815). According to Sweitzer and Volkwein (2009), the first model is "the resource/reputation model" (p. 815). The proponents of this model (Astin, 1985) accentuate the importance of institutional resources, which constitute input indicators. Thus, "the resource/reputation model" considers financial resources, external funding, faculty quality, student quality indicators, and peer ratings as proxies for quality in higher education.

The second model, in Sweitzer and Volkwein's (2009) opinion, is "the client-centered model" (p. 815). The advocates of this model (Seymour, 1992) view students, parents, and college employees as "customers" and associate quality with "student and

alumni satisfaction, faculty availability and attentiveness, and robust student services" (Sweitzer & Volkwein, 2009, p. 815).

The third model, as Sweitzer and Volkwein's (2009) suggested, is "the strategic investment model" (p. 815). The supporters of this model (Burke, 2005; Ewell, 1994) prioritize return on investment and cost-benefit analysis and link productivity measures, such as expenditure per student, graduation rates, time-to-degree, and admission yields, to quality of higher education institutions.

The last model on Sweitzer and Volkwein's (2009) list is "the talent development model" (p. 815) suggested first by Astin in 1985. The tenets of Astin's (1985) model posit that the primary goal of higher education institutions is the multifaceted development of students and faculty, which includes the development of "knowledge, skills, attitudes, and interests" (p. 815-816). Consequently, the quality of institutions should be gauged according to their ability to provide the maximum opportunities for their students and faculty to realize their full potential.

Consistent with the debate about the quality of higher education institutions in general, Martins (2005) argued that "the products of b-schools are credence goods, the quality of which constituents cannot evaluate easily" (p. 708). Therefore, the problems with the operationalization of a definition of business school quality underlie the issues of "construct validity" (Cronbach & Meehl, 1955) of business school rankings. Cronbach and Meehl (1955) contended that "construct validity must be investigated whenever no criterion or universe of content is accepted as entirely adequate to define the quality to be measured" (p. 282). In their methodologies, rankings of business schools rely on proxies

for quality, and "the proxies for educational quality turn out to be flimsy at best" (Gladwell, 2011, p. 72).

The problem of construct validity of business school rankings has attracted strong criticism from a number of researchers. Gioia and Corley (2002) characterized the rankings criteria as "a ham-fisted attempt to quantify 'quality'" (p. 111). Martins (2005) referred to the rankings criteria as "poor indicators of quality" (p. 702). DeAngelo et al. (2005) asserted that the rankings criteria are "too narrow to capture any real quality differences that might exist among business schools" (p. 4). Dichev (1999) claimed that "the rankings are mechanical and incomplete aggregations of 'noisy' information" (p. 212). Gioia and Corley (2002) also expressed concern about the weights assigned to rankings criteria, which add to the problem of construct validity.

Despite frequent lamentations concerning the methodological inaccuracies in business school rankings, the indicators of quality that are used by the five major ranking systems can be aligned with models of quality in higher education proposed by Sweitzer and Volkwein (2009). *U.S. News & World Report* employs quality indicators that constitute the core of "the resource/reputation model" (Sweitzer & Volkwein, 2009, p. 815). In the 2012 edition of *U.S. News & World Report Best Graduate Schools* guidebook, the full-time M.B.A. school rankings methodology (Morse & Flanigan, 2011) lays great emphasis on business school reputation that is gauged based on the results of two reputational surveys, which form 40% of the overall rank. The first survey elicits the opinion of business school deans and directors of accredited master's programs in business, and its results constitute 25% of the overall rank. The second survey is

administered to corporate recruiters, and the recruiter assessment score makes up 15% of the overall rank. The next part of the rankings – placement success (35%) – gauges the outcome measures of the business school quality. The average starting salary and bonus and employment rates for graduates of a full-time M.B.A. program constitute 14% and 21%, respectively. Finally, the student selectivity component (25%) represents the input measures of the business school quality. The student selectivity component includes mean GMAT scores (16.25%), mean undergraduate GPA (7.5%), and acceptance rate (1.25%).

Among the five most influential business school rankings[5], *Bloomberg Businessweek* rankings of full-time M.B.A. programs (*Bloomberg Businessweek*, 2011) relies most heavily on the reputational component. The rankings are based on three surveys of M.B.A. graduates, three surveys of M.B.A. recruiters, which were conducted in 2010, 2008, and 2006, and a review of faculty research publications. The rank based on surveys of M.B.A. graduates and the rank based on surveys of M.B.A. recruiters contribute 45% each to the final ranking score. The heavy emphasis on reputation in the *Bloomberg Businessweek* and *U.S. News & World Report* rankings incur scholars' criticism because it tends "to favour the best-known universities in many nations" (Marginson & van der Wende, 2007, p. 312). The "anchoring effect" (Bowman & Bastedo, 2011, p. 431) and the ability of well-known university "brands" "to generate "halo" effects" (Marginson & van der Wende, 2007, p. 320) can skew the results of

---

[5] Appendix A presents the results for the top 30 schools of business in each of the five ranking systems in 2010 and 2011.

reputational surveys and turn them into "popularity contests" (Altbach, 2006, p. 2). In the *Bloomberg Businessweek* rankings, the rank based on the review of faculty research constitutes the final component and includes publications in 20 top academic journals and faculty books reviewed in *The New York Times, The Wall Street Journal, and Bloomberg Businessweek* over a five-year period. The faculty research component contributes 10% to the final ranking.

*The Financial Times* Global M.B.A. rankings, which map the global position of business schools from different countries, also rely on the outcome and input quality indicators from "the resource/reputation model" (Sweitzer & Volkwein, 2009, p. 815). *The Financial Times* Global M.B.A. rankings create the list of the 100 top M.B.A. programs based on the analysis of three main areas that include alumni salaries and career development, the diversity and internationalization of the business school and its M.B.A. program, and research productivity of each school. Among the 20 criteria that *The Financial Times* Global M.B.A. rankings 2011 (*The Financial Times*, 2011) employ in determining each school's overall rank, weighted salary (US$) and salary percentage increase are weighted the heaviest (20% each), thus making the outcome quality measures the most influential ones. The second most heavily weighted criterion is *FT* research rank (10%), which is determined based on faculty research published in 45 academic and practitioner journals. The remaining 50% of *The Financial Times* Global M.B.A. rankings is composed of value for money rank, careers progress rank, aims achieved rank, placement success rank, percentage of recent graduates employed at three months, alumni recommend rank, percentage of women faculty, percentage of women

students, percentage of women on board, percentage of international faculty, percentage of international students, percentage of international members on board, international mobility rank, international experience rank, languages, percentage of faculty with doctorates, and *FT* doctoral rank[6].

*The Economist* Global full-time M.B.A. rankings and ranking breakdowns by region (e.g., Europe ranking, North America ranking, and Asia & Australasia ranking) are based on quality indicators both from "the resource/reputation model" and from "the client-centered model" (Sweitzer & Volkwein, 2009, p. 815). *The Economist 2011* Global full-time M.B.A. rankings (*The Economist*, 2012a) produce a list of the global top 100 business schools based on two web-based questionnaires, one for business schools and one for students and recent graduates. In *The Economist's* methodology, student and alumni ratings make up 20% of the total rankings, and 80% of the rankings is based on data provided by business schools (*The Economist*, 2012b). The 13 rankings criteria are grouped into four categories, such as open new career opportunities (35%), personal development/educational experience (35%), increase in salary (20%), and potential to

---

[6] In *The Financial Times* rankings, measures of the diversity of staff, board members, and students by gender and nationality and measures of the quality of faculty do not show the percentages per se. Business schools are ranked against each other by calculating a *z*-score for each criterion. For the three gender-related criteria, for example, schools with a 50:50 (male/female) composition receive the highest possible score. *Z*-scores for each criterion are weighted as outlined in the methodology and added together to produce a final score for a business school (*The Financial Times*, 2015).

network (10%). The "open new career opportunities" category includes such measures as diversity of recruiters, assessment of career services, jobs found through the career services, and student assessment. The "personal development/educational experience" category consists of four measures, such as faculty quality, student quality, student diversity, and education experience. The "increase in salary" category consists of two measures, such as the increase in salary after graduation and leaving salary. Finally, the "potential to network" category is comprised of such measures as breadth of alumni network, internationalism of alumni, and alumni effectiveness. Overall, the four categories consist of 13 measures of quality. Each of the 13 measures of quality is assessed through corresponding indicators, which total 21. Educational experience, for example, is evaluated based on four indicators of quality, such as student rating of program content and range of electives, range of overseas exchange programs, number of languages on offer, and student assessment of facilities and other services. As a result, multiple indicators of each measure of quality allow a more comprehensive assessment of the quality of an M.B.A. program.

In contrast to the four M.B.A. rankings described above, *Forbes* Best Business Schools rankings are grounded in "the strategic investment model" of quality in higher education proposed by Sweitzer and Volkwein (2009) and present a cost-benefit analysis of full-time M.B.A. programs. The Best Business Schools rankings 2011 (Badenhausen, 2011) are based on the calculation of the return on investment received by the graduates from the class of 2006. *Forbes* rankings of M.B.A. programs compare the salary earned by graduates over their first five years after receiving their M.B.A. to their opportunity

cost that consists of foregone earnings, tuition, and required fees. *Forbes* rankings of M.B.A. programs adjust earnings for cost of living expenses and discount tuition to account for differences between in-state and out-of-state tuition and the nonrefundable financial aid.

Although all five business school rankings methodologies are subject to a litany of criticism, Bradshaw (2007) pointed out the advantages of "this proliferation" (p. 57) in the number of rankings. First, as Bradshaw (2007) indicated, different business school rankings "rank on different criteria and therefore paint a broader picture of the strengths of various programmes" (p. 57). Second, Bradshaw (2007) thought that "To some degree the various rankings validate each other" (p. 57) because the similar criteria, which different rankings use, "produce similar rankings" (p. 57). As a result, the differences in business school rankings methodologies allow evaluating schools from different perspectives, and the similarities in methodologies validate their quality.

In line with this view, in the field of legal education, where the *U.S. News & World Report Best Graduate Schools* guidebook holds a monopoly over rankings, scholars (Sauder & Espeland, 2006; Stake, 2006) advocate the need for a multiplicity of rankings, which might serve students, the public, and law schools more effectively. In this respect, Sauder and Espeland (2006) referred to the example of business school rankings and asserted that "By comparing the effects of law school rankings (a field where one ranker dominates) to those of business schools (a field where multiple influential rankers evaluate schools), we are able to suggest how the effects of law school

rankings might be mitigated if more than one credible ranking system existed for legal education" (p. 206).

**Critique of Methodologies**

The studies that criticize the methodologies employed in the business school rankings are diverse and numerous, which is also true about the studies of law school rankings, although the latter tend to be less methodologically complex (e.g., Klein & Hamilton, 1998; Schmalbeck, 1998; Stake, 2006). In general, it is possible to classify the studies of business school rankings methodologies into four groups based on similarities in their conceptual framework.

**Measuring the reputation of business schools.** The first group of studies is grounded in the fact that "rankings are significant drivers of a school's reputation" (Peters, 2007, p. 49). As a result, these studies (Rindova et al., 2005; Sweitzer & Volkwein, 2009) bring the reputational component of business school rankings into focus in their analysis. In their study of 49 schools of business ranked as best business schools by *U.S. News & World Report*, Sweitzer and Volkwein (2009), for example, used the average peer assessment rating for each school as a proxy for assessing prestige among professional schools. Consistent with the resource/reputation model of quality in higher education (Sweitzer & Volkwein, 2009), the researchers' conceptual framework included the predictor variables drawn both from the *U.S. News & World Report* rankings methodology and prior research on academic prestige. As a result, Sweitzer and Volkwein (2009) used institutional characteristics (full-time enrollment, full-time faculty, and non-resident tuition), faculty resources (student-faculty ratio), student resources

(average GMAT scores, acceptance rate, and average undergraduate GPA), faculty outcomes (publication per faculty 2001-2005), and student outcomes measures (starting salary of graduates and percent of graduates employed at graduation) as major predictors of business school prestige. Sweitzer and Volkwein's (2009) analysis demonstrated that "full-time enrollments, average GMAT scores, and the starting salaries of graduates explain almost 88% of the variance in the *USNWR* reputation scores" (p. 828).

In a similar vein, Rindova, Williamson, Petkova, and Sever (2005) attempted to explore the "antecedents" (p. 1033) of business school reputation and their relationship with media rankings indicators of quality. In their study, Rindova et al. (2005) integrated the economics and institutional perspectives of the reputation study and suggested that they represent "two distinct dimensions of reputation" (p. 1034). Rindova et al. (2005) asserted that the economics perspective emphasizes "the *perceived quality* dimension of organizational reputation" (p. 1034), which shows how stakeholders assess a certain organizational attribute. The institutional perspective, in Rindova, Williamson, Petkova, and Sever's (2005) view, accentuates "the *prominence* dimension of organizational reputation" (p. 1034) that represents "the collective awareness and recognition that an organization accumulated in its organizational field" (p. 1034). In accordance with their conceptual framework, Rindova et al. (2005) used prominence, perceived quality, and price premium as measures of reputation. Researchers measured prominence by the number of recruiters who nominated a given school that was obtained from an online survey of 1,600 corporate recruiters administered by Harris Interactive in 2000. Rindova et al. (2005) gauged perceived quality as the average of recruiters' ratings of a given

school on 13 attributes of student quality. Finally, researchers used the mean starting base salary, which *U.S. World & News Report* published in 2001, as a measure of price premium. As independent variables, the researchers used quality of inputs, as measured by the average GMAT scores, quality of productive assets, as gauged by the average number of years of academic experience of a business school's faculty, and media rankings, as represented by *Businessweek's* 1998 rankings. Rindova et al. (2005) included two more independent variables in their model, which are not commonly used in other scholars' models. The first independent variable is expert intermediary certifications, as measured by the total number of publications of the faculty of a given school over five years. The second independent variable is the affiliation with high-status actors, as represented by the prestige of doctorate degrees possessed by a school's faculty. Rindova et al. (2005) used the Gourman Report (Gourman, 1997) "to assign a Ph.D. degree-prestige score" (p. 1041) to each professor affiliated with a given school and then averaged the individual faculty scores.

As a result of the measured variable path analysis in LISREL 8.53 (Jöreskog & Sörbom, 2001), Rindova et al. (2005) concluded that "Media rankings had the largest significant effect on prominence (-.51), as well as a significant total effect on price premium (-.30)" (p. 1043). Overall, Rindova, Williamson, Petkova, and Sever's (2005) results suggested that "prominence, which derives from the choices of influential third parties vis-à-vis an organization, contributes significantly to the price premium associated with having a favorable reputation" (p. 1033).

A more recent study by Safón (2009) revisited the problem of identifying the predictors of the reputation of top business schools in the United States. Safón (2009) operationalized reputation as perceived by three important stakeholders: recruiters of M.B.A. graduates, business school deans and directors of M.B.A. programs, and prospective M.B.A. students. Employing path analysis of the Multiple Indicators and Multiple Causes (MIMIC) model, Safón (2009) examined the impact of the six predictors of business school reputation, namely, research performance, quality of students, program value, placement success, the order in which M.B.A. programs entered the market, and media rankings of business schools. Out of the six antecedents, only three were shown to make a statistically significant impact on business school reputation: quality of students $\beta = .38$, $p < .00$, program value $\beta = .11$, $p = .01$, and media rankings $\beta = .54$, $p < .00$ (Safón, 2009, p. 218). In general, the results of Safón's (2009) study demonstrated that "reputation of a business school is a unidimensional construct, implying homogeneity in the perceptions of the stakeholders analyzed, and fundamentally hinges on two antecedents: quality of students and position in the business school media rankings" (p. 204).

**Predictors of business school rankings.** The second group of studies (Liebowitz, 2003; Morgeson & Nahrgang, 2008; Safón, 2007; Schatz & Crummer, 1993; Siemens, Burton, Jensen, & Mendoza, 2005) focuses primarily on the predictors of media rankings. Liebowitz (2003) and Siemens, Burton, Jensen, and Mendoza (2005) found a strong correlation between the media rankings and research productivity. Liebowitz (2003) explored the impact of research productivity in terms of citation counts and

publications on the overall rank in the three major ranking systems – *Businessweek, U.S. News & World Report,* and *The Financial Times* in 1999. Both the *Businessweek* and *The Financial Times* rankings of business schools have a specific research component in their methodology (faculty research rank based on the number of faculty publications), whereas *U.S. News & World Report* lacks this component. For this reason, Liebowitz (2003) used academic reputation as a proxy for the research prestige of a business school in the *U.S. News & World Report* rankings. The results of Liebowitz's (2003) study demonstrated that the overall rank was highly correlated with the research rank in all three ranking systems, with .625, .895, and .793 Pearson correlations for *Businessweek*, *U.S. News & World Report*, and *The Financial Times*, respectively (p. 19). The magazine research rankings were shown to be also correlated with one another.

Siemens, Burton, Jensen, and Mendoza's (2005) study provided a more in-depth analysis of the relationship between research productivity and the overall ranking score. Siemens et al. (2005) revealed that "research productivity from 1986-1998 is a significant predictor ($t = 5.58$; $P$ [*sic*] $< .001$) and explains more than 40% of the variance in the 2000 MBA program ratings" (p. 473) in *U.S. News & World Report*. The researchers also conducted a longitudinal analysis and examined the Spearman rank order correlations between research productivity measures over time (1986-1989, 1990-1993, 1994-1998, and the total time frame (1986-1998)) and the *U.S. News & World Report* M.B.A. program rankings in 2000 and 1995-1999 and the *Businessweek* M.B.A. rankings in 2000. Siemens, Burton, Jensen, and Mendoza's (2005) findings demonstrated that the Spearman rank order correlations between research productivity and M.B.A. rankings

were relatively strong across the time periods, ranging from .47 to .67 (p. 474). At the same time, the Spearman rank order correlations for the *Businessweek* M.B.A. rankings in 2000 were very similar to those for the *U.S. News & World Report* M.B.A. rankings in 2000 for both the overall period (1986-1998) ($r_s$ = .65 and $r_s$ = .62) and the three separate 1986-1989 ($r_s$ = .65 and $r_s$ = .63), 1990-1993 ($r_s$ = .67 and $r_s$ = .64), and 1994-1998 ($r_s$ = .60 and $r_s$ = .56) time frame periods (Siemens et al., 2005, p. 474).

The diverse system of the predictors of M.B.A. media rankings allows for various approaches to their study. Along with the impact of research productivity on the overall rank score, other researchers (Schatz & Crummer 1993; Safón, 2007) focus their attention on different predictors of media rankings. Schatz and Crummer (1993), for example, showed that there was a high correlation between the rank of a business school and certain input-output measures of its M.B.A. program, such as GMAT scores and starting salary. Schatz and Crummer (1993) demonstrated that the rankings created using a simple average of GMAT scores and starting salary produced the same results as the *Businessweek* and the *U.S. News & World Report* M.B.A. rankings. In a similar vein, Safón (2007) hypothesized that student characteristics, such as GPA and GMAT scores, program value, as measured by the starting salary, and research performance were primary predictors of M.B.A. rankings. The specific attribute of Safón's (2007) analysis lies in the fact that the researcher explored four major ranking systems (*Businessweek*, *U.S. News & World Report*, *The Wall Street Journal*, and *Forbes*) in "a single construct" (p. 223) contrary to other authors who concentrated predominantly on the *Businessweek* and the *U.S. News & World Report* ranking systems. As a result of this analysis, Safón

(2007) showed that student characteristics, as measured by GPA and GMAT scores, and program value added, as measured by the starting salary, proved to be "worthy predictors" (p. 225) as the regression of the combined media rankings on these predictors produced "a high adjusted $R^2$ (.76, $F = 52.18$, $p < .001$, maximum VIF = 2.34)" (p. 225). At the same time, the research productivity was shown to be not significant ($\Delta R^2 = .01$, $p > .05$) as a predictor of the M.B.A. ranking score (Safón, 2007, p. 225).

Morgeson and Nahrgang (2008) conducted a more comprehensive study of the predictors of business school rankings. While exploring the *Businessweek* M.B.A. rankings from 1990 to 2004, Morgeson and Nahrgang (2008) expanded their theoretical framework to include six predictors of these rankings. First, Morgeson and Nahrgang (2008) assumed that "the length of time the business school had been in existence" (p. 30), as measured by the year the school was founded, impacted its rankings. Except the year of foundation, Morgeson and Nahrgang (2008) also added the 1988 *Businessweek* rankings as an important school characteristic that defines its prestige. Second, Morgeson and Nahrgang (2008) hypothesized that the quality of students, as measured by the average GMAT score and work experience of the entering class in a particular year, was "likely to be important for MBA program success" (p. 30). Third, Morgeson and Nahrgang (2008) theorized that placement success, as measured by the average starting salary of the graduating class, could be "a critical aspect of any MBA program" (p. 30). Finally, stemming from previous research, Morgeson and Nahrgang (2008) suggested that faculty research productivity is an important factor in predicting the rankings. Morgeson and Nahrgang's (2008) results demonstrated that "the six variables accounted

for between 64% and 100% of the variance in the *Businessweek* rankings" (p. 31). The two school characteristic variables, the 1988 *Businessweek* rankings and the year the school was founded, when viewed together, "accounted for between 37% and 81% of the variance in the ratings in any particular year" (p. 31). Morgeson and Nahrgang (2008) also revealed that student characteristics accounted for between 0% and 63% of unique variance in the rankings; placement characteristics accounted for between 1% and 54%; and finally, research characteristics accounted for between 0% and 38% (p. 33). As a result, Morgeson and Nahrgang (2008) concluded that student characteristics were demonstrated to be the most important predictors of rankings, followed by placement characteristics, and research characteristics appeared to be relatively unimportant in accounting for variance in rankings, which validated the results of the previous studies of M.B.A. rankings.

**Measurement error and the problem of reliability of business school rankings.** The third group of studies (Dichev, 1999; Fee, Hadlock, & Pierce, 2005; Holbrook, 2007; Morgeson & Nahrgang, 2008) focuses on the measurement error in the rankings of business schools. In this respect, Dichev (1999) analyzed "first-order autocorrelation in changes as an empirical proxy for predictability" (p. 203) of the *Businessweek* and the *U.S. News & World Report* M.B.A. rankings. Dichev's (1999) study rested on the premise in statistics (Dodge, 2010; Montgomery, Jennings, & Kulanchi, 2008) that positive autocorrelation indicates that positive departures from the mean tend to be followed by positive departures from the mean, and negative departures from the mean tend to be followed by negative departures from the mean. By contrast,

negative autocorrelation is characterized by a propensity for positive departures to follow negative departures, and vice versa (Dodge, 2010; Montgomery et al., 2008). Based on this statistical assumption, Dichev's (1999) first hypothesis about the predictability of M.B.A. rankings stated that "a positive autocorrelation in the rankings' changes might be interpreted as evidence of "momentum" or "feedback" effects in business schools' real performance, while a negative autocorrelation might be indicative of "fight back" effects in real performance" (p. 202). Based on the analysis of the *Businessweek* and the *U.S. News & World Report* M.B.A. rankings in 1988, 1990, 1992, 1994, and 1996, Dichev (1999) found that first-order autocorrelation was statistically significant and negative, ranging from -.27 for the *U.S. News & World Report* to -.46 for the *Businessweek* rankings (p. 204). At the same time, the *U.S. News & World Report* rankings had a lower *SD* in changes (1.97) as compared to the *Businessweek* rankings (3.84) (Dichev, 1999, p. 204). These results suggested that "a large portion of the rankings changes predictably reverses in the future" (Dichev, 1999, p. 202). Moreover, the comparison of the two ranking systems showed that the higher negative autocorrelation for the *Businessweek* rankings indicated "more transitory changes" (Dichev, 1999, p. 205), and the higher *SD* indicated "a larger average magnitude of changes" (Dichev, 1999, p. 205) in the *Businessweek* rankings.

The second hypothesis in Dichev's (1999) study tested "the timeliness of the rankings" (p. 202). Dichev (1999) assumed that if changes in business school rankings are "a timely reflection of real changes in the underlying unobservable "business school quality" (p. 202), the concurrent *Businessweek* and *U.S. News & World Report* M.B.A.

rankings should be strongly positively correlated. A weak or no-positive correlation between the two concurrent rankings, in Dichev's (1999) view, suggested that "the two rankings use unrelated information" (p. 202). The results of the statistical tests for correlation between short-run (1990-92, 1992-94, and 1994-96) and long-run (1990-96) concurrent rankings changes for the *Businessweek* and the *U.S. News & World Report* rankings revealed that the cross-magazine correlation is close to zero (.01 and .23, respectively) and statistically insignificant (*p*-value of .968 and .366, respectively) (Dichev, 1999, p. 206). As a result, Dichev (1999) concluded that "the two rankings are based on completely different information sets" (p. 207) and suggested a possible explanation that "the lack of comovement in the rankings could be due to the fact that one or both of the rankings react to new information with a lag" (p. 207). Overall, from Dichev's (1999) standpoint, the reversibility and consequent unreliability of the two rankings as predictors of business school quality may be explained by three major reasons, such as "magazine-specific changes in methodology" (p. 207), "fight back" effects" (p. 208) in schools' performance, and "the mechanical incorporation of noisy information" (p. 208). Thus, the major flaw of the *Businessweek* and the *U.S. News & World Report* rankings stems from the measurement error in the rankings methodology.

Fee, Hadlock, and Pierce (2005) corroborated and extended Dichev's (1999) evidence on the reversibility of the *Businessweek* and the *U.S. News & World Report* rankings and cross-correlation between the two ranking systems. Using the rankings data from the 1990-2002 editions, which is a significantly larger sample than in Dichev's (1999) study, Fee et al. (2005) determined that both the *Businessweek* and the *U.S. News*

*& World Report* rankings demonstrate strong negative first-order correlation (-.38 and -.20, respectively), which is statistically significant (*p*-value of .000 and .008, respectively) (p. 152). These results allowed Fee et al. (2005) to conclude that the *Businessweek* and the *U.S. News & World Report* rankings changes are "partially predictable" (p. 151).

While testing the correlation structure of the *Businessweek* and the *U.S. News & World Report* rankings series, Fee et al. (2005) suggested two approaches. First, Fee et al. (2005) matched each *Businessweek* rankings with the *U.S. News & World Report* rankings and obtained a high statistically significant cross-magazine correlation (.832) (p. 152), which contradicted Dichev's (1999) results. Nevertheless, the issues of these magazines cannot be matched directly as the *Businessweek* rankings are released biannually, and the *U.S. News & World Report* rankings are released every year. Therefore, to overcome this measurement problem, Fee et al. (2005) proposed to consider the changes in the two rankings over two-year periods. As a result, the authors obtained a low cross-magazine correlation (.007), which was not statistically significant (*p*-value of .949) (p. 152), and confirmed Dichev's (1999) findings of "very little correlation" (Fee et al., 2005, p. 151) in the updated *Businessweek* and *U.S. News & World Report* rankings series.

While Dichev's (1999) and Fee, Hadlock, and Pierce's (2005) studies questioned the idea of rankings' predictability and statistically supported the finding that M.B.A. rankings are reversible, Morgeson and Nahrgang (2008) claimed that business school rankings are "highly stable over time" (p. 26). Researchers based their study on the data

from the biennial *Businessweek* rankings from 1988 to 2004. *Businessweek* ranked the top-20 business schools from 1988–1994, the top-25 from 1996–1998, and the top-30 from 2000–2004, which constituted the sample size for the study. Morgeson and Nahrgang (2008) suggested three different ways "to quantify the stability of these rankings" (p. 29). First, Morgeson and Nahrgang (2008) examined the magnitude of the rank order correlations between ranking years and found out that the correlations ranged from .66 to .92, with an average correlation of .82 (p. 29). Thus, Morgeson and Nahrgang (2008) concluded that "on average, 67% of the variance in the current year ranking is shared with the previous year's ranking" (p. 29). Second, researchers explored the turnover rate of the rankings and revealed that, on average, 1.63 business schools dropped off the *Businessweek* ranking list in any given year, and the maximum dropouts never exceeded four schools (Morgeson & Nahrgang, 2008, p. 30). As a result, "across years, the rankings contain 84% to 97% of the same schools" (Morgeson & Nahrgang, 2008, p. 30). Finally, Morgeson and Nahrgang (2008) identified the schools that had always been ranked. Researchers determined that of the 20 business schools, which were on the original 1988 *Businessweek's* list of the top business schools, 18 schools had always been ranked. Moreover, Morgeson and Nahrgang (2008) revealed that only 35 business schools had ever been ranked by *Businessweek* that testifies to "the considerable stability in the rankings" (p. 30). Assuming that the 20 schools from the original 1988 *Businessweek's* list are likely to remain on future ranking lists, which now include 30 schools, Morgeson and Nahrgang (2008) pointed out that "schools hoping to become

ranked are competing against each other for only 10 places, not 30 places" (p. 30), which demonstrates the intensity of the ranking competition.

Morgeson and Nahrgang's (2008) findings about the stability of the *Businessweek* rankings attested to the face validity of the methodology of business school rankings, which McGuire (1995) suggested first with respect to the *U.S. News & World Report* college rankings methodology almost two decades ago. Face validity, according to Bornstein's (2004) definition, refers to "an estimate of the degree of which a measure is clearly and unambiguously tapping the construct it purports to assess" (Face Validity, para.1). McGuire (1995) claimed that "The face validity of "America's Best Colleges" is probably its strong suit" (p. 57). Sweitzer and Volkwein (2009) developed this point by arguing that "People tend to have a preconceived notion of which schools are listed at the top each year, and indeed, most of those schools are at the top of the *USNWR* lists each year" (p. 814). Based on this premise, the stability of business school rankings provides face validity to these rankings and lends some credence to the very idea of educational rankings amidst its strong criticism.

The measurement error as one of the most widely criticized flaws of the business school rankings methodologies gained a new perspective in Holbrook's (2007) study, which explored the accuracy of the weighting scheme of the *U.S. News & World Report* M.B.A. rankings. Unlike other studies of rankings of educational institutions, which demonstrated the impact of changing weights (McGuire, 1995, for example), Holbrook (2007) focused his attention on the weights assigned to the two most debatable subjective indicators of the quality of business schools – the peer assessment score and the recruiter

assessment score. The researcher regressed each of these scores on the eight objective indicators of quality, namely, GPA, GMAT score, acceptance rate, average starting salary and bonus, percentage of graduates employed at graduation, percentage of graduates employed three months after graduation, out-of-state tuition and fees, and total full-time enrollment, in order to obtain "(1) the best possible prediction or explanation of the subjective evaluations using the objective measures (PredPeer and PredRecruiter) and (2) estimates of possible distorting biases due to neglected considerations giving rise to the residual error terms in each regression (ResidPeer and ResidRecruiter)" (Holbrook, 2007, p. 5). As Holbrook (2007) explained it, the residual error terms "capture the effects of omitted variables that cause subjective impressions to differ from those expected on the basis of available objective measures" (p. 5-6). At the same time, the possible predicted scores of peer and recruiter assessment represent "our best guesses as to what the subjective evaluations would be in the absence of possible biases due to neglected considerations" (Holbrook, 2007, p. 6). By substituting the peer and recruiter assessment scores with the standardized predicted scores of peer and recruiter assessment in the *U.S. News & World Report's* original formula, Holbrook (2007) obtained a "Corrected Score" (p. 6), which "indicates our best estimate of what the Computed Score would be if the subjective peer and recruiter ratings reflected objective characteristics in a manner that left them free from potential biases introduced by extraneous factors" (p. 6). Based on the corrected score, Holbrook (2007) produced an alternative ranking of the top 20 business schools ranked in the 2005 *U.S. News & World Report*. In order to accentuate the differences between the original and corrected ranks, Holbrook (2007) calculated "the

Adjusted Impact on Ranking" scores (p. 8) that showed the extent to which some business schools suffered from unfavorable biases and some benefited from favorable ones "due to neglected considerations" (p. 9). Based on the assumption that subjective biases "result from obvious omitted variables" (p. 9), Holbrook (2007) proposed "to include these influences in the model so as to improve its predictive and explanatory power" (p. 9). Thus, instead of criticizing business school rankings for the obvious measurement error, Holbrook (2007) suggested the avenue to their improvement by exploring "what missing or unmeasured influences account for the discrepancies found in the original versus the corrected rankings" (p. 9).

**Redefining methodologies of business school rankings.** While the majority of researchers concentrate on criticizing the statistical validity, the weighting scheme, and the methodologies employed in the business school rankings, some scholars attempt to redefine the rankings methodologies. Tracy and Waldfogel (1997), for example, suggested "a market-based methodology" (p. 1) for evaluating the performance of M.B.A. programs that sought "to distinguish the quality of a program from the quality of its students" (p. 1). Tracy and Waldfogel (1997) created a complex statistical algorithm to determine a program's value added that excluded subjective indicators of quality gauged by the *Businessweek* and the *U.S. News & World Report* rankings series. The novelty of Tracy and Waldfogel's (1997) approach also lay in the fact that they viewed value added as "measured by the graduates' salaries, after accounting for student characteristics and job attributes" (p. 1). First, researchers adjusted starting salaries of the M.B.A. graduates by region, occupation, and industry in order to make salary data

"comparable across programs" (Tracy & Waldfogel, 1997, p. 5). Second, Tracy and Waldfogel (1997) standardized the adjusted salary for "the quality of inputs" (p. 7), such as GMAT scores, percent of students with at least one year of full-time work experience, undergraduate GPA, percent of students with a graduate degree, and the ratio of acceptances to applications, and weighted these student characteristics into "a single index of student quality" (p. 7). Researchers regressed the adjusted starting salary on average student characteristics and obtained a measure of "a school's "value added" as "the residual from this regression" (Tracy & Waldfogel, 1997, p. 7). At the same time, the measure of "student quality" was "the fitted value from this salary regression" (Tracy & Waldfogel, 1997, p. 7). Tracy and Waldfogel (1997) produced an alternative ranking of a sample of business schools based on the measure of value added and compared it to the 1991 *Businessweek* and *U.S. News & World Report* rankings of business programs. In the alternative rankings, M.B.A. programs which were ranked high on value added were not ranked high in both the *Businessweek* and the *U.S. News & World Report* rankings because these programs did not enjoy high reputation and did not attract top quality students.

Another attribute of Tracy and Waldfogel's (1997) study consists in their attempt to identify the flaws in the *Businessweek* and the *U.S. News & World Report* rankings. For this purpose, researchers regressed the *Businessweek's* top 20 and the *U.S. News & World Report's* top 25 business schools on each program's adjusted salary, student quality, and value added. The results demonstrated that there was a strong correlation between the *Businessweek* and the *U.S. News & World Report* rankings and the adjusted

salary rankings. By contrast, when the adjusted salary ranking was disaggregated into student quality and value added, the results indicated that the *Businessweek* and the *U.S. News & World Report* rankings attach "a much larger weight to student quality than to value added" (Tracy & Waldfogel, 1997, p. 18).

As researchers realize the value of the market-based approach for prospective students in selecting M.B.A. programs, they seek new ways of evaluating the quality of business schools from the market perspective. In this respect, Fisher, Kiang, and Fisher (2007) proposed to extend Tracy and Waldfogel's (1997) model and measured both value added and best value of programs. The novelty value of Fisher, Kiang, and Fisher's (2007) study also lies in the fact that the researchers used the data envelopment analysis (DEA) technique, which was developed by Charnes, Cooper, and Rhodes (1978) "to assess the relative efficiencies of multi-input production units" (Cook & Seiford, 2009, p. 1). The DEA methodology is based on "the benchmark of economic efficiency, Pareto optimality" (Brownstein, 1980, p. 93). Pareto optimality postulates that "a given decision making unit (DMU) is inefficient if some other DMU, or some combination of other DMUs, can produce at least the same amounts of all outputs with less of some resource input and not more of any other resource; conversely, a DMU is said to be Pareto efficient if the above is not possible" (Lewin, Morey, & Cook, 1982, p. 403).

Since its advent in 1978, the DEA methodology has evolved into various complicated models (Cook & Seiford, 2009) and has become "an increasingly popular management tool" (Fisher et al., 2007, p. 73). Grounded in mathematical theory, a DEA model compares the operating unit under evaluation with other units based on multiple

input and output measures and generates an efficiency score that ranges from 0 to 1. Researchers (Colbert, Levary, & Shaner, 2000; Fisher et al., 2007) argued that an efficiency value of 1 indicates that an operating unit is the most efficient. Correspondingly, an efficiency value smaller than 1 shows the degree of relative efficiency.

The DEA methodology as an efficiency measurement tool has been widely applied in both the public and private sectors. In the field of higher education in the United States, a number of research studies have investigated the relative efficiency of academic departments (Ahn, 1987; Tauer, Fried, & Fry, 2007), public research universities (Zheng & Stewart, 2000), both public and private doctoral-degree-granting higher education institutions (Ahn & Seiford, 1993; Blalark, 2012), and liberal arts colleges (Eckles, 2010). With respect to the evaluation of M.B.A. programs in the U.S., only a few studies used the DEA method (Colbert et al., 2000; Fisher et al., 2007; Haksever & Muragishi, 1998; Hsu, Chao, & James, 2009).

Haksever and Muragishi (1998), for example, measured the efficiency of the top 20 business schools identified by the *Businessweek* rankings and compared them to the 20 less prestigious business schools. The results of the study indicated that there was no "significant difference between the top 20 (88.2%) and second 20 (87.6%) in terms of overall average ($t = 0.208$, $p$-value = 0.836)" efficiency (Haksever & Muragishi, 1998, Discussion, para 6). Despite general perceptions, Haksever and Muragishi (1998) also found that private business schools do not outperform their public counterparts in terms of overall average efficiency.

Unlike Colbert, Levary, and Shaner's (2000) study of the relative efficiency of M.B.A. programs that used subjective measures of output, such as student and recruiter satisfaction with different aspects of the program, Fisher, Kiang, and Fisher's (2007) study, which was mentioned earlier, used the DEA methodology to analyze value added from the market perspective. In order to gauge value added, Fisher et al. (2007) included GMAT and GPA scores as input measures and starting salaries and bonuses, employment rates at graduation, and employment rates three months after graduation as output measures. In order to calculate best value of programs, researchers used tuition, cost of living, and length of a program as additional input attributes. As a result, researchers avoided the pitfalls of subjective bias in data analysis. Fisher et al. (2007) produced the DEA efficiency rankings and compared them with the 2004 *Businessweek* and the 2005 *U.S. News & World Report* rankings. The study demonstrated that DEA's value added and best value rankings were different from the media rankings. *U.S. News & World Report* ranked University of Taxes A&M, for example, as 32, and *Businessweek* did not even include it in the second tier, whereas DEA ranked this business school 1st both in terms of value added and best value. As a result, Fisher et al. (2007) argued that because the DEA methodology "overcomes the deficiency introduced by subjective weights" (p. 75), the rankings generated by the DEA can be helpful and "provide students with unbiased guidelines for selection" (p. 75) of the M.B.A. program.

The most recent study of the efficiency of M.B.A. programs (Hsu et al., 2009) revisited the issue of an efficiency gap between top and nontop business schools. Hsu, Chao, and James (2009) compared pure technical efficiency (PTE), which indicates how

efficiently a business school uses its resources, scale efficiency (SE), which demonstrates how productive the scale size is, and technical and scale efficiency (TSE) for the more expensive 10 elite U.S. M.B.A. programs and their less expensive 10 nontop counterparts. Researchers found that the 10 elite programs outperformed the 10 nonelite programs in terms of scale efficiency and technical and scale efficiency. Nevertheless, the 10 nonelite programs achieved parity with the elites in terms of pure technical efficiency. Based on these findings, Hsu et al. (2009) recommended that prospective M.B.A. applicants conduct a cost and benefit analysis and avoid decision-making based on the media rankings or DEA's analysis. The specific value of Hsu, Chao, and James's (2009) study also lies in the fact that researchers revealed the sources of inefficiency by analyzing the input and output measures of the less efficient business schools, which may provide administrators with avenues to schools' improvement.

**Rationale for Present Research**

The review of previous research on rankings of business schools has demonstrated that the proliferation of media rankings has generated "a parallel proliferation of academic commentary" (Peters, 2007, p. 50). The studies of business school rankings fall into two main groups. The first group of studies presents both an opinionated judgment and empirical evidence of the impact of rankings on different aspects of business schools' behavior. The second group of studies criticizes the methodologies that are employed in rankings and questions the statistical validity of their results.

Given the overwhelming complexity of the substance of rankings and a variety of scholars' opinions, I identify the following six lacunas in the study of M.B.A. rankings in the United States, each of which provides the rationale for the present study. First, existing studies of the rankings of M.B.A. programs focus primarily on the *Businessweek* and the *U.S. News & World Report* rankings, leaving *Forbes*, *The Economist*, and *The Financial Times* rankings out of the limelight of scholars' attention. Nevertheless, *Forbes*, *The Economist*, and *The Financial Times* rankings of M.B.A. programs, which are published by the prestigious business media, have a strong potential for influencing the business community. The prospective M.B.A. students who typically gain some work experience in the industry before applying to a business school differ significantly from traditional college applicants. They are more likely to be influenced by the standards and values imposed by the business community, which *Forbes*, *The Economist*, and *The Financial Times* rankings exemplify. Moreover, *Forbes*, *The Economist*, and *The Financial Times* rankings evaluate the quality of business schools from a completely different perspective than the *Businessweek* and the *U.S. News & World Report* rankings. *Forbes* rankings, for example, give the estimate of value added in terms of return on investment for the M.B.A. degree, which many scholars attempt to gauge in various studies. *The Economist* rankings include personal development, educational experience, and assessment of career services as important measures of a business school's quality, which provides some insight into the quality of the teaching and learning process. Finally, *The Financial Times* rankings assess U.S. business schools' quality in the competitive global environment. As a result, the comparative analysis of the

methodologies of all five existing ranking systems of full-time M.B.A. programs in the U.S. is warranted and might yield useful information on their validity.

Second, the majority of recent studies concentrate predominantly on the most prestigious business schools, which constitute the first tier in the ranking systems. Only a few studies include business schools from the second tier in their analysis. This research gap may be explained by the fact that not all media rankings of M.B.A. programs, the *Businessweek* rankings, for example, provide information about business schools beyond the second tier. Nevertheless, *Forbes* and *The Financial Times* include at least 100 business schools in their rankings, and the *U.S. News & World Report* rankings provide information about all M.B.A. programs in the United States, even those which are not accredited by the Accreditation Council for Business Schools and Programs (ACBSP)[7]. Therefore, the inadequate sample size, which is biased towards the most prestigious business schools, may undermine the statistical validity of the previous research on business school rankings.

Third, the small body of studies that explore business school rankings from the longitudinal perspective were based on the data collected from the maximum of nine

---

[7] "ACBSP accredits business, accounting, and business-related programs at the associate, baccalaureate, master, and doctorate degree levels worldwide. Recognized by the Council for Higher Education Accreditation (CHEA) in 2001 and again in 2011, ACBSP was the first to offer specialized business accreditation at all degree levels" (The Accreditation Council for Business Schools and Programs (ACBSP), 2013, About ACBSP. Discover ACBSP, para. 2).

issues of media rankings reports (for example, Dichev, 1999; Morgeson & Nahrgang, 2008). Taking into consideration the fact that rankings methodologies evolved over time, the longitudinal analysis of all issues of rankings reports since their inception might allow determining whether their validity has improved as well.

Fourth, no study explored the correlation between the ranking of schools of business and the overall ranking of the institutions in which they are situated. The results of this analysis might shed light on the impact of business schools on the overall performance of higher education institutions in the competitive educational arena.

Fifth, despite the efforts of the international community of education professionals, researchers, and policy-makers to develop a framework for the assessment of the quality of ranking systems, only a few studies (for example, Stolz et al., 2010) used the Berlin Principles on Ranking of Higher Education Institutions (IREG, 2006) to assess the quality of rankings of higher education institutions. With respect to business school rankings, no study examined the quality of M.B.A. program rankings through the lens of the Berlin Principles on Ranking of Higher Education Institutions. Therefore, the "IREG Observatory" (IREG, 2011) methodology for assessing ranking systems based on the Berlin Principles on Ranking of Higher Education Institutions might provide a new perspective on the study of rankings of business schools.

Finally, the obvious controversy in some studies of business school rankings methodologies, such as the findings about stability versus reversibility in the rankings, for example, necessitates revisiting contentious issues. The longitudinal study that draws on all five existing rankings of full-time M.B.A. programs in the U.S.A. might yield

different results regarding their validity and quality, which might reshape prospective students' perspectives, business schools' policies, and how recruiters use rankings in hiring decisions.

# CHAPTER III: METHODOLOGY

In concordance with the research framework, this chapter consists of four sections that describe the methodology utilized to analyze the four research questions in this study. The first section details the procedures used to assess the full-time M.B.A. programs through the lens of the Berlin Principles on Ranking of Higher Education Institutions. The second section describes the methods employed to assess the reliability and validity of the full-time M.B.A. program rankings. Specifically, the second section elaborates on the methods of estimating predictability of changes of the rankings, estimating stability of the rankings, and assessing construct validity of the rankings with CFA of multitrait-multimethod matrices. The third section specifies the procedures utilized to examine the relationship between rankings and reputation of business schools and rankings and reputation of higher education institutions. Finally, the fourth section explains the methods used to explore how rankings of full-time M.B.A. programs affect funding streams to institutions of higher education.

## Research Question 1

**Ranking of rankings: Assessment of the full-time M.B.A. program rankings through the lens of the Berlin Principles on Ranking of Higher Education Institutions.** This part of the research addresses the first research question and assesses the quality of the five full-time M.B.A. ranking systems through the lens of the Berlin Principles on Ranking of Higher Education Institutions. For this assessment, I adopted the methodology suggested by "IREG Observatory" (IREG, 2011) that is based on the Berlin Principles on Ranking of Higher Education Institutions (See Appendix B).

Although the IREG Ranking Audit methodology is grounded on the Berlin Principles on Ranking of Higher Education Institutions for the most part, it incorporates some new criteria that do not pertain directly to the Berlin Principles. As the IREG team explained, the Berlin principles were "a first attempt to define general principles of good ranking practice" (IREG, 2011, p. 6). Therefore, the IREG, building on the Berlin Principles, elaborated a methodology that covered a broad spectrum of aspects of the quality of rankings. The IREG Ranking Audit criteria refer to five dimensions of rankings: first, criteria on purposes, target groups, and their basic approach; second, criteria on the methodology; third, criteria on the publication and presentation of results; fourth, criteria on transparency and responsiveness; and last, criteria on quality assurance (IREG, 2011). Table 1 presents the short description of the criteria that refer to each of the five dimensions of rankings.

Criteria are assessed with numerical scores on a scale from 1 to 6, with 1 being *not sufficient/not existing*, 2 – *marginally applied*, 3 – *adequate*, 4 – *good*, 5 – *strong*, and 6 – *distinguished*. The IREG team proposed to divide the 20 criteria into 10 core criteria with a weight of two, and 10 regular criteria with a weight of one, as "not all criteria are of the same relevance" (IREG, 2011, p. 8). The numerical score for each criterion is multiplied by its weight based on the attribution of the criterion to the core or regular group. The weighted scores for all the criteria are added together to get the overall score for the ranking system, with the maximum total score of 180. Table 1 provides the details of how to reach the maximum score.

"In order to establish the IREG Ranking audit as a quality label" (IREG, 2011,

Table 1

*IREG Ranking Audit Criteria*

| Criterion | Weight |
|---|---|
| **Purpose, Target Groups, Basic Approach** | |
| 1. The purpose of the ranking and the (main) target groups should be made explicit | 2 |
| 2. Rankings should recognize the diversity of institutions | 2 |
| 3. Rankings should specify the linguistic, cultural, economic, and historical contexts of the educational systems being ranked | 1 |
| **Methodology** | |
| 4. Rankings should choose indicators according to their relevance and validity | 2 |
| 5. The concept of quality of higher education institutions is multidimensional and multi-perspective (…). Good ranking practice would be to combine the different perspectives | 1 |
| 6. Rankings should measure outcomes in preference to inputs wherever possible | 1 |
| 7. Rankings have to be transparent regarding the methodology used for creating the rankings | 2 |
| 8. If rankings are using composite indicators the weights of the individual indicators have to be published. Changes in weights over time should be limited and due to methodological or conception-related considerations | 2 |
| 9. Data used in the rankings must be obtained from authorized, audited and veritable data sources and/or collected with proper procedures for professional data collection | 2 |
| 10. The basic methodology should be kept stable as much as possible | 1 |
| **Publication and Presentation of Results** | |
| 11. The publication of a ranking has to be made available to users throughout the year either by print publications and/or by an online version of the ranking | 1 |
| 12. The publication has to deliver a description of the methods and indicators used in the ranking | 1 |

(continued)

Table 1 (continued)

| Criterion | Weight |
|---|---|
| 13. The publication of the ranking must provide scores of each individual indicator used to calculate a composite indicator in order to allow users to verify the calculation of ranking results | 2 |
| 14. Rankings should allow users to have some opportunity to make their own decision about the relevance and weights of indicators | 1 |
| **Transparency, Responsiveness** | |
| 15. Rankings should be compiled in a way that eliminates or reduces errors | 1 |
| 16. Rankings have to be responsive to higher education institutions included/participating in the ranking | 2 |
| 17. Rankings have to provide a contact address in their publication (print, online version) | 1 |
| **Quality Assurance** | |
| 18. Rankings have to apply measures of quality assurance to ranking processes themselves | 2 |
| 19. Rankings have to document their internal processes of quality assurance | 1 |
| 20. Rankings should apply organizational measures that enhance the credibility of rankings | 2 |
| **Maximum Total Score** (with 6-point scale of assessment) | **180** |

*Note.* From *The IREG Ranking Audit Manual* (p. 9), by The International Ranking Expert Group, 2011. Retrieved from http://www.ireg-observatory.org/pdf/ranking_audith_audit.pdf

p. 8), the IREG methodology specifies the threshold of 60% of the maximum total score, which produces the minimum total score of 108, to approve the ranking system. Because the present study did not aim to approve or disapprove ranking systems, the minimum total score of 108 was not used as a threshold parameter in the assessment of the quality of rankings of business schools.

Unlike the IREG Ranking Audit assessment, which only establishes the threshold for a positive audit decision at 60% (the score of 108) of the maximum total score of 180 and neither compares the institutions' scores nor publishes the audit scores, this study assessed the five full-time M.B.A. rankings in the United States according to the IREG criteria and ranked them based on the total scores.

In accordance with the IREG recommendations[8], three raters rated the five full-time M.B.A. rankings based on the IREG methodology specified above. Additionally, IREG suggests that "Members of an Audit Team should represent a range of professional experience in higher education, quality assurance and the assessment of higher education institutions or systems" (IREG, 2011, p. 12). Therefore, the first rater was the author of the study, and the other two raters represented two business schools, one from the Midwest and the other from the East Coast. As a result, the team of raters consisted of professionals who have expertise in both higher education as a field and, more specifically, in graduate management education.

The three raters were provided with the most recent results of the five ranking systems of full-time M.B.A. programs in the U.S.A., namely, *U.S. News & World Report* (2013)*, Bloomberg Businessweek* (2012), *The Financial Times* (2013), *The Economist* (2013), and *Forbes* (2013). In order to increase the reliability of the raters' ratings, several precautions were taken to ensure that the raters consistently used the same information to rate the five full-time M.B.A. rankings. First, the raters were provided

---

[8] IREG recommends that the Audit Team consist of three to five members (IREG, 2011, p. 12).

with a package of rankings materials that included: 1) information about each of the five ranking systems under analysis, 2) the most recent results of the five ranking systems of full-time M.B.A. programs, and 3) the methodology used to produce the results. Second, each package of rankings materials was supplemented with guidelines about the type of information from the package that was needed to rate rankings according to each of the IREG criteria. Finally, each package of rankings materials contained references to the five rankings websites and the most recent publications in order to provide insight into the workings of rankings of full-time M.B.A. programs in the U.S.A. and the needs of their target audience.

Because the scoring procedure involves multiple raters, "it is important to estimate the degree of interrater reliability, as this value has important implications for the validity of the study results" (Stemler, 2004, para. 3). Given the plethora of reliability coefficients (Popping (1988), for example, compared 43 reliability coefficients only for nominal data), this study drew on the classification of interrater reliability measures proposed by Stemler (2004). Stemler (2004) suggested classifying interrater reliability measures into three categories: 1) consensus estimates, 2) consistency estimates, and 3) measurement estimates. Based on Stemler's (2004) premise that the appropriate approach to estimating interrater reliability should depend on both the purpose of the study and the data available, I selected consensus estimates of interrater reliability as the best fit for this study.

According to Stemler (2004), consensus or agreement estimates are based on the assumption that raters can come to "exact agreement about how to apply the various

levels of a scoring rubric to the observed behaviors" (Consensus estimates: General description, para. 1). Moreover, agreement estimates allow calculating the summary scores "by averaging the scores given by all of the judges since high interrater reliability indicates that the judges agree about how to apply the rating scale" (Stemler, 2004, Consensus estimates: General description, para. 4). Krippendorff (2004b) and Hayes and Krippendorff (2007) provided a detailed comparison of seven different agreement coefficients. Krippendorff's (2004a, 2004b) α (alpha) suits both the data and the purpose of this study. Krippendorff's α (alpha) can "handle multiple coders; nominal, ordinal, interval, ratio, and other metrics; missing data; and small sample sizes" (Krippendorff, 2004b, p. 428). Therefore, Krippendorff's α (alpha) seems to be a proper estimate of the interrater reliability for the design of this study that involves three raters and is based on ordinal[9] data (a 6-point scale) and a small sample size consisting of the 20 audit criteria

[9] As the analysis did not involve any statistical procedure that required that data satisfy the assumption of normality, but involved the arithmetic procedures that are necessary to calculate the average score for each of the IREG Ranking Audit criterion and the total score for each ranking system (e.g., addition, multiplication, and division), ordinal data (a 6-point scale) were treated as interval data. This is in line with some researchers' assumption that "If the categories are viewed as equal-appearing intervals... then the ordinal measure could be treated as interval/ratio for statistical purposes" (Nardi, 2006, p. 56). Nevertheless, for the statistical procedures used in educational research, other researchers (Harwell & Gatti, 2001) recommend that ordinal data be rescaled to interval data.

of the five full-time M.B.A. programs in the U.S.A. Moreover, as an agreement estimate of interrater reliability, Krippendorff's α (alpha) allows calculating the average score for each of the IREG Ranking Audit criterion that contributes to the maximum total score obtained for each of the five rankings.

As IREG does not specify the method used to estimate interrater reliability in the IREG Ranking Audit Manual (IREG, 2011), this study utilized Krippendorff's (Krippendorff, 2004a, 2004b) α (alpha) as a measure of interrater agreement. Krippendorff (2011) defined the general form of the alpha coefficient as follows:

$$\alpha = 1 - \frac{D_o}{D_c}$$

where "$D_o$ is the observed disagreement among values assigned to units of analysis and $D_c$ is the disagreement one would expect when the coding of units is attributable to chance rather than to properties of these units" (Krippendorff, 2011, p. 1). Computations of Krippendorff's α (alpha) can be performed by hand (see the procedure in Krippendorff, 2011) and with a number of statistical software packages, such as SPSS and SAS (code available in Hayes & Krippendorff, 2007 or at http://www.afhayes.com), AgreeStat (code available in Gwet, 2011), and R (code available at http://cran.r-project.org/web/packages/irr/irr.pdf).

In terms of the minimum acceptable level of the agreement estimate of interrater reliability, Krippendorff (2004b) recommended $\alpha \geq .800$. In cases where "tentative conclusions" (Krippendorff, 2004a, p. 241) were made, Krippendorff (2004a) proposed that $\alpha \geq .667$ be acceptable. Based on these recommendations, I accepted $\alpha \geq .667$ as the minimum level of agreement below which the average scores obtained from the three

raters for each criterion were to be rejected as too unreliable. If the minimum level of agreement was not achieved for some particular criteria of rankings, the raters were asked to discuss the points on which they were in disagreement in order to reach consensus on how to score those criteria.

After calculating the overall score for each of the five ranking systems of full-time M.B.A. programs based on the raters' ratings, I rank ordered the ranking systems to produce the ranking of rankings of M.B.A. programs. Furthermore, I compared the resulting ranking of M.B.A. rankings, which is based on the objective quality assessment of these ranking systems by higher education professionals, with students' evaluation of the importance of various M.B.A. rankings for their decision to enroll as reported in the 2012 mba.com Prospective Students Survey (Graduate Management Admission Council, 2012b). The results from the Prospective Students Survey are presented in Figure 1.

Based on the results in Figure 1, the prospective students' ranking of rankings includes *The Financial Times* (1 rank), *U.S. News & World Report* (2 rank), *Bloomberg Businessweek* (2 rank), *Forbes* (4 rank), and *The Economist* (5 rank). In order to compare the results of ranking of rankings, the Spearman rank order correlation coefficient (Kutner, Nachtshein, Neter, & Li, 2005) was used. The comparison of the results of the two rankings of rankings of M.B.A. programs allowed identifying the differences in rankings assessment between the two groups of stakeholders: a group of experts inside academe and a group of potential customers outside academe.

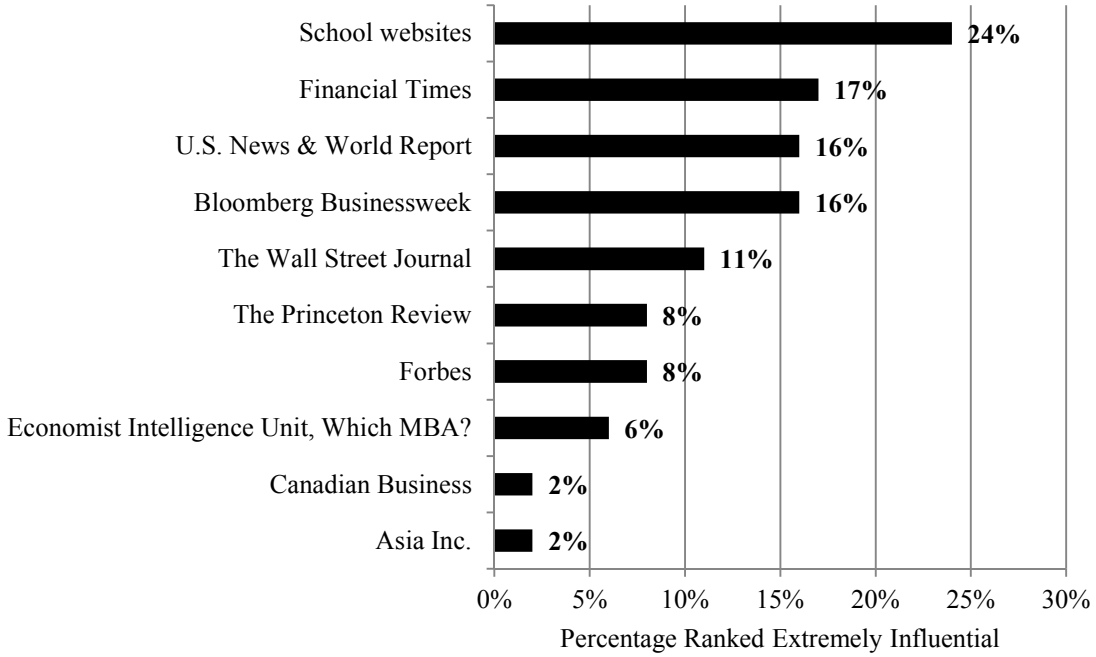


*Figure 1.* Influence of a school's website on the decision to apply compared with ranking publications. From *The mba.com Prospective Students Survey* (p. 27), by Graduate Management Admission Council, 2012. Retrieved from http://www.gmac.com/~/media/Files/gmac/Research/prospective-student-data/2012prospectivestudentssr.pdf. Copyright 2012 by Graduate Management Admission Council.

**Research Question 2**

**Reliability of the full-time M.B.A. rankings in the United States: Estimating predictability of changes of the rankings.** This part of the research addresses the second research question and estimates reliability of the four full-time M.B.A. ranking systems by gauging the predictability of changes in the rankings. The underlying rationale for this research question lies in the fact that predictability of rankings may be

of paramount importance for prospective M.B.A. students. As Dichev (1999) noted, "Forecasting ability is likely a key consideration to business school applicants because there is a 2–3-year lag between the ranking that determines the applicants' school choice and the ranking that determines the employers' demand for a given school's graduates" (p. 202).

In general, as summarized by Webb, Shavelson, and Haertel (2006), classical test theory's reliability coefficients fall into three major groups: coefficients of stability, equivalent forms reliability coefficients, and internal consistency reliability coefficients. Although predictability of changes in data observations refer to time series analysis, and coefficients of stability were developed in the framework of classical test theory, both measures present the two facets of the reliability construct and seek "to estimate the consistency of scores across repeated observations" (Webb, Shavelson, & Haertel, 2006, p. 81).

The present study builds on Dichev's (1999) proposition that first-order autocorrelation of error terms[10] can serve as "an empirical proxy for predictability" (p. 203) of rankings of M.B.A. programs. Dichev's (1999) proposition is based on the postulation in time series analysis that the error terms in one period are correlated with the error terms in another period (Sheather, 2009; Weiss, 2010). As a result, if the data

---

[10] "The autocorrelation of lag $l$ is the correlation between $Y$ and values of $Y$ lagged by $l$ periods, i.e., between $Y_t$ and $Y_{t-l}$, i.e.,

$$\text{Autocorrelation}(l) = \sum_{t=l+1}^{n}(y_t - \bar{y})(y_{t-l} - \bar{y}) / \sum_{t=l+1}^{n}(y_t - \bar{y})^2$$"

(Sheather, 2009, p. 308).

set exhibits positive autocorrelation (or serial correlation), the positive departures from the mean in one time series will result in the proportionate positive departures from the mean in the other time series, and the negative departures from the mean in one time series will result in the proportionate negative departures from the mean in the other time series (Weiss, 2010). By contrast, negative autocorrelation indicates that an increase happened in one time series will result in a proportionate decrease in the other time series, and vice versa (Weiss, 2010). Thus, when time series data show significant autocorrelation, it is possible to forecast whether the future rankings will demonstrate stability of either an upward or downward trend or will be completely unpredictable.

Based on this forecasting technique, this study seeks to revisit Dichev's (1999) findings about unpredictability of the *U.S. News & World Report* and the *Bloomberg Businessweek* ranking systems and expand them by including *The Financial Times* and *Forbes* rankings into statistical analysis. In his study, Dichev (1999) found statistically significant ($p = .001$) negative first-order autocorrelation (-.46) for the *Bloomberg Businessweek* rankings and statistically significant ($p < .000$) negative first-order autocorrelation for *U.S. News & World Report* in two cases: using only observations from schools that have rankings for all years (-.27) and using all available observations (-.34) (p. 204). The inclusion of the four ranking systems of full-time M.B.A. programs into analysis allowed comparison of the rankings with regard to the reliability of the information that they provide for prospective students and the general public. The

*Economist* ranking system was not included in the analysis because of limited data availability[11].

The other aspect of the present research that is different from Dichev's study is the data used for the analysis. While Dichev's study was based on the top 20 business schools ranked by *Bloomberg Businessweek* in 1988, 1990, 1992, 1994, and 1996, and the top 25 business schools ranked by *U.S. News & World Report* in 1990, 1991, 1992, and 1993 and top 50 business schools ranked in 1994, 1995, 1996, 1997, and 1998, this study analyzed the data from the four full-time M.B.A. rankings over the 14-year period. The present study hypothesized that methodological changes made to the M.B.A. ranking

---

[11] Typically, publishers of M.B.A. program rankings provide a full list of ranked programs with a detailed description of their methodology on their websites. They provide a shorter list of the top-ranked M.B.A. programs in their paper publications. *U.S. News & World Report*, for example, provides information about all ranked business schools (around 180) in the United States on its website. At the same time, *U.S. News & World Report* publishes a list of about 60 top business schools in its paper edition. Publishers of rankings update information about rankings on their websites annually (or biannually, depending on the rankings publisher) and typically do not provide archived information about previous rankings. *The Economist* provides very limited information on rankings of business schools (only the list of top schools without rankings criteria) in its paper edition. Therefore, I could not obtain information about rankings criteria for *The Economist* ranking system over the 14-year period to construct a longitudinal data set.

systems between 2000 and 2013 might yield positive results with regard to the reliability of the rankings, which would suggest that rankings are responsive to criticism and have made an effort to improve their methodologies. Table 2 presents a description of the data from the four rankings of full-time M.B.A. programs used to test for first-order autocorrelation. Table 3 presents the number of observations for each of the four rankings across years. Each ranking system has its own panel structure depending on the

Table 2

*Data Sources for Estimating Predictability of Changes of the Rankings*

| Year | *U.S. News & World Report* | *Bloomberg Businessweek* | *The Financial Times* | *Forbes* |
|---|---|---|---|---|
| 2000 | 50 schools | 30 schools | 45 schools | - |
| 2001 | 53 schools | - | 51 schools | old methodology |
| 2002 | 50 schools | 30 schools | 56 schools | - |
| 2003 | 50 schools | - | 56 schools | 50 schools |
| 2004 | 50 schools | 30 schools | 57 schools | - |
| 2005 | 50 schools | - | 58 schools | 60 schools |
| 2006 | 51 schools | 30 schools | 57 schools | - |
| 2007 | 50 schools | - | 61 schools | 56 schools |
| 2008 | 50 schools | 30 schools | 58 schools | - |
| 2009 | 50 schools | - | 57 schools | 75 schools |
| 2010 | 51 schools | 57 schools | 56 schools | - |
| 2011 | 72 schools | - | 53 schools | 74 schools |
| 2012 | 69 schools | 63 schools | 52 schools | - |
| 2013 | 72 schools | - | 50 schools | 70 schools |

Table 3

*Number of Observations Used for Estimating Predictability of Changes of the Rankings*

| Ranking System | Number of rankings publications over 14-year period | Balanced panel structure | Unbalanced panel structure |
|---|---|---|---|
| *U.S. News & World Report* | 14 | 546 | 768 |
| *Bloomberg Businessweek* | 7 | 175 | 270 |
| *The Financial Times* | 14 | 392 | 765 |
| *Forbes* | 6 | 264 | 385 |

number of times the rankings were calculated between 2000 and 2013[12] and the number of indicators it uses to calculate the rank. Table 3 presents the number of observations for both balanced and unbalanced panel structures. Balanced panel data consist of observations from schools that have rankings for all years. Unbalanced data include all available observations for all schools.

The *Bloomberg Businessweek* and *Forbes* rankings data sets did not have missing values. *The Financial Times* rankings data set had a small number of missing values in the *FT* doctoral rank, international experience rank, and percentage of faculty with doctorates indicators. The total number of data entries in *The Financial Times* rankings balanced data set was 8,232, with only 42 missing values for seven business schools. Out

[12] *U.S. News & World Report* and *The Financial Times* rank full-time M.B.A. programs every year, while *Bloomberg Businessweek* and *Forbes* produce rankings of full-time M.B.A. programs every other year.

of 16,065 data entries in the unbalanced data set, 120 were missing values for 19 business schools. Because the number of business schools in the longitudinal balanced data set for *The Financial Times* rankings equaled 28, exclusion of seven business schools with missing values would significantly decrease the sample size. "Unconditional mean imputation" (Fox, 2016, p. 611) was used to replace the missing values with the mean of the observed data for the rankings indicator in both balanced and unbalanced data sets. In a few cases in the unbalanced data set where a business school had observations only for one out of 14 years and mean imputation was impossible, the missing values were replaced with the neighboring nonmissing values.

The *U.S. News & World Report* rankings had four missing values in the acceptance rate indicator for two business schools in both balanced and unbalanced data sets. The missing values were imputed with means.

The publishers of rankings did not provide any explanation for missing values. Several reasons could account for missing values. First, business schools could refuse to provide information for a particular indicator in the rankings. Second, business schools could provide insufficient or inadequate information that prevented the publishers from calculating a score for a particular rankings indicator. Third, missing data could occur due to human error on the part of the publishers of rankings while processing the data received from business schools. Examination of the data did not reveal any pattern in the missing data. Both high- and low-ranked business schools had missing values. Different business schools had missing values for different years. Therefore, the data were

assumed to be missing completely at random, which means that missing values did not introduce substantial bias.

Depending on the number of indicators used in the ranking systems to calculate the rank, four generalized multiple regression models when the random error terms follow a first-order autoregressive process (AR1)[13] with continuous AR(1) correlation structure were fit to balanced and unbalanced data sets described in Table 3. Because the nature of the time series trend was not of interest in this study, and the best fitting models (depending on the covariance structure) were different for different ranking systems, in order to ensure consistency, the autoregressive (AR1) covariance model was chosen as a special case of the more general ARMA (autoregressive moving average) covariance models.

The generalized multiple regression model when the random error terms follow a first-order autoregressive process (AR1) with continuous AR(1) correlation for the *U.S.*

---

[13] The generalized multiple regression model when the random error terms follow a first-order autoregressive process is:

$$Y_t = \beta_0 + \beta_1 X_{t1} + \beta_2 X_{t2} + \ldots + \beta_{p-1} X_{t,p-1} + \varepsilon_t$$

$$\varepsilon_t = \rho \varepsilon_{t-1} + u_t$$

where:

$\rho$ is a parameter such that $|\rho| < 1$

$u_t$ are independent $N(0, \sigma^2)$. (Kutner, Nachtshein, Neter, & Li, 2005, pp. 484-485).

The parameter $\rho$ is called *the autocorrelation parameter*.

*News & World Report* ranking system is:

$$Y_t = \beta_0 + \beta_1 X_{t1} + \beta_2 X_{t2} + ... + \beta_9 X_{t9} + \varepsilon_t$$

$$\varepsilon_t = \rho\varepsilon_{t-1} + u_t$$

where:

$\beta_0, \beta_1, ..., \beta_9$ are parameters

$X_{t1}, X_{t2}, ..., X_{t9}$ represent the year of rankings edition (i.e., time as linear effect) and the eight rankings criteria in the *U.S. News & World Report* rankings at time *t* as predictor variables

$t = 1, ..., 14$ (14 years when rankings were published between 2000 and 2013, see Tables 2 and 3).

The generalized multiple regression model when the random error terms follow a first-order autoregressive process (AR1) with continuous AR(1) correlation for the *Bloomberg Businessweek* ranking system is:

$$Y_t = \beta_0 + \beta_1 X_{t1} + \beta_2 X_{t2} + \beta_3 X_{t3} + \beta_4 X_{t4} + \varepsilon_t$$

$$\varepsilon_t = \rho\varepsilon_{t-1} + u_t$$

where:

$\beta_0, \beta_1, ..., \beta_4$ are parameters

$X_{t1}, X_{t2}, X_{t3}$, and $X_{t4}$ represent the year of rankings edition (i.e., time as linear effect) and the three rankings criteria in the *Bloomberg Businessweek* rankings at time *t* as predictor variables

$t = 1, ..., 7$ (seven years when rankings were published between 2000 and 2013, see

Tables 2 and 3).

The generalized multiple regression model when the random error terms follow a first-order autoregressive process (AR1) with continuous AR(1) correlation for *The Financial Times* ranking system is:

$$Y_t = \beta_0 + \beta_1 X_{t1} + \beta_2 X_{t2} + ... + \beta_{20} X_{t20} + \varepsilon_t$$

$$\varepsilon_t = \rho\varepsilon_{t-1} + u_t$$

where:

$\beta_0, \beta_1, ..., \beta_{20}$ are parameters

$X_{t1}, X_{t2}, ..., X_{t20}$ represent the year of rankings edition (i.e., time as linear effect) and the 19 rankings criteria in *The Financial Times* rankings at time *t* as predictor variables

$t = 1, ..., 14$ (14 years when rankings were published between 2000 and 2013, see Tables 2 and 3).

The generalized multiple regression model when the random error terms follow a first-order autoregressive process (AR1) with continuous AR(1) correlation for the *Forbes* ranking system is:

$$Y_t = \beta_0 + \beta_1 X_{t1} + \beta_2 X_{t2} + \beta_3 X_{t3} + \beta_4 X_{t,4} + \varepsilon_t$$

$$\varepsilon_t = \rho\varepsilon_{t-1} + u_t$$

where:

$\beta_0, \beta_1, ..., \beta_4$ are parameters

$X_{t1}, X_{t2}, X_{t3}$, and $X_{t4}$ represent the year of rankings edition (i.e., time as linear effect) and the three rankings criteria in the *Forbes* rankings at time *t* as predictor variables

$t = 1, ..., 6$ (six years when rankings were published between 2000 and 2013, see Tables 2

and 3).

To test for first-order autocorrelation of the error terms, this study used the Ljung-Box test (Ljung & Box, 1978) that is a modified version of the Box-Pierce test (Box & Pierce, 1970) for lack of fit in time series models, such as autoregressive (AR), moving average (MA), and autoregressive-integrated moving average (ARIMA) models. In both tests, "the test statistics are (approximate) $\chi^2$ statistics based on estimates of the autocorrelations up to order *p*" (Kleiber & Zeileis, 2008, p. 105). According to Box and Pierce (1970), the Box-Pierce statistic is *n* times the sum of squared autocorrelations of residuals

$$Q(r\hat{}) = n\sum_{k=1}^{m} r\hat{}_k^2$$

where *n* is the sample size, *r(k)* is the autocorrelation at lag *k*, and *m* is the number of lags being tested. The Ljung and Box (1978) modified test weighs the squared autocorrelation at lag *k* by $(n + 2)/(n - k)$ ($k = 1, ..., m$). The Ljung-Box test statistic is computed as

$$Q(r\hat{}) = n(n+2)\sum_{k=1}^{m}(n-k)^{-1} r\hat{}_k^2$$

(Ljung & Box, 1978, p. 298).

The Ljung-Box test has the null hypothesis that the autocorrelation of the error terms or disturbances is 0 or the model does not exhibit lack of fit. The alternative hypothesis states that the autocorrelation of the error terms or disturbances is not equal to 0 or the model exhibits lack of fit. Computation of the Ljung-Box test statistic was performed with the R version 3.1.0 (2014-04-10) statistical software (in package "lmtest"). The acf (autocorrelation function) in R was used to calculate the magnitude of the autocorrelation of the error terms and to produce a plot of the autocorrelations with an

indication of their statistical significance (Sheather, 2009).

**Reliability of the full-time M.B.A. rankings in the United States: Estimating reliability as stability of the ranking scores.** Another way to estimate reliability concerning stability of the five full-time M.B.A. ranking systems is grounded in measurement theory. In the domain of classical test theory, researchers (Cortina, 1993; Rodriguez & Maeda, 2006; Thorndike & Thorndike-Christ, 2010; Webb, Shavelson, & Haertel, 2006) delineate three major ways to assess reliability depending on the three primary sources of error: stability estimates of reliability, equivalence estimates of reliability, and internal consistency estimates of reliability. If "variation of the individual from time to time, as well as variation due to the operation of measurement" (Thorndike & Thorndike-Christ, 2010, p. 123), is the focus of the study, the correlation between scores at different occasions produces "a test-retest reliability – a coefficient of stability" (Webb et al., p. 85).

Following this premise, this study utilized the Spearman rank order correlations across 14 years as an estimate of stability of the five full-time M.B.A. ranking systems. As *U.S. News & World Report* stopped dividing M.B.A. programs into tiers/quartiles in 2002, *Bloomberg Businessweek* showed schools in the second tier (without ranking them) only in 2008, and *The Financial Times*, *The Economist*, and *Forbes* rankings never stratified business schools in their publications, this study estimated the stability of each of the five M.B.A. rankings as a total set, without dividing them into tiers. By estimating the magnitude of the Spearman rank order correlations across years as a means of assessing stability and first-order autocorrelations of error terms as a means of assessing

predictability of rankings, this study sought to triangulate its results with regard to reliability of the full-time M.B.A. ranking systems. Table 4 presents a description of the data from the five rankings of M.B.A. programs used to test for the Spearman rank order correlations across years[14]. In order to maximize the sample size and portray a more accurate picture of the rankings, the unbalanced data sets for the five full-time M.B.A. ranking systems were used. Unbalanced data included all available observations for all schools across years. "Available-case analysis" (Fox, 2016, p. 610) or pairwise deletion of missing values was used to compute the statistic. The R version 3.1.0 (2014-04-10) statistical software was used to calculate the Spearman rank order correlations across years.

**Construct validity of the full-time M.B.A. rankings in the United States: Estimating construct validity with CFA of multitrait-multimethod matrices.** In the absence of a consensus among experts on how to define quality in higher education (Brooks, 2005; Usher & Savino, 2006), rankings of the M.B.A. programs, which are viewed as a proxy for the quality of business schools, face the problem of construct

---

[14] With longitudinal data, "Correlations between trial *j* and *k* were calculated by using all subjects that had both observations at times *j* and *k*" (Weiss, 2010, p. 58). Based on the rankings data analyzed in this study, the five correlation matrices for the five rankings had different numbers of trials or times at which observations were made and different numbers of lags between observations. The correlation matrix for the *U.S. News & World Report* rankings, for example, had 14 times at which observations were made, while the correlation matrix for *The Economist* rankings had only six times.

Table 4

*Data Sources for Estimating Reliability as Stability of the Ranking Scores*

| Year | *U.S. News & World Report* | *Bloomberg Businessweek* | *The Financial Times* | *The Economist* | *Forbes* |
|---|---|---|---|---|---|
| 2000 | 50 schools | 30 schools | 45 schools | - | - |
| 2001 | 53 schools | - | 51 schools | started in 2002 | old methodology |
| 2002 | 50 schools | 30 schools | 56 schools | data unavailable | - |
| 2003 | 50 schools | - | 56 schools | data unavailable | 50 schools |
| 2004 | 50 schools | 30 schools | 57 schools | data unavailable | - |
| 2005 | 50 schools | - | 58 schools | data unavailable | 60 schools |
| 2006 | 51 schools | 30 schools | 57 schools | data unavailable | - |
| 2007 | 50 schools | - | 61 schools | 47 schools | 56 schools |
| 2008 | 50 schools | 30 schools | 58 schools | 47 schools | - |
| 2009 | 50 schools | - | 57 schools | data unavailable | 75 schools |
| 2010 | 51 schools | 57 schools | 56 schools | 46 schools | - |
| 2011 | 72 schools | - | 53 schools | 48 schools | 74 schools |
| 2012 | 69 schools | 63 schools | 52 schools | 48 schools | - |
| 2013 | 72 schools | - | 50 schools | 52 schools | 70 schools |

validation. As Cronbach and Meehl (1955) posited, "Construct validity must be investigated whenever no criterion or universe of content is accepted as entirely adequate to define the quality to be measured" (p. 282). As a result, it is necessary to develop evidence for the validity of measures that are used in the rankings of M.B.A. programs to

determine the hypothetical construct of quality of an M.B.A. program. Campbell and Fiske (1959) argued that two types of evidence are required "for the establishment of construct validity, *discriminant* validation as well as convergent [*sic*] validation" (p. 81). Convergent validity is demonstrated by associations among measures of the same construct obtained from "independent measurement procedures" (Campbell & Fiske, 1959, p. 81). Discriminant validity "requires that a new construct be substantially less correlated with measures of conceptually unrelated constructs than with other indicators of that construct" (Strauss & Smith, 2009, p. 6), as measured by the same method.

In order to examine both convergent and discriminant validity evidence systematically, Campbell and Fiske (1959) suggested the multitrait-multimethod matrix (MTMM) approach. Although developed within the context of assessing validity of individual traits in psychology, the multitrait-multimethod matrix (MTMM) approach applies as well to assessment of validity of measures (or traits) of programs and organizations (Bagozzi & Phillips, 1982; Bagozzi, Yi, & Phillips, 1991). Campbell and Fiske (1959) outlined the four desiderata or requirements for evaluating a multitrait-multimethod correlation matrix. First, correlations for the same trait measured with different methods, which serve as evidence of convergent validity, "should be significantly different from zero and sufficiently large to encourage further examination of validity" (Campbell & Fiske, 1959, p. 82). Second, "a validity value for a variable should be higher than the correlations obtained between that variable and any other variable having neither trait nor method in common" (Campbell & Fiske, 1959, p. 82). Third, correlations of the same trait measured with different methods should be higher

than correlations among different traits measured with the same method. Finally, it is required that "the same pattern of trait interrelationship be shown in all of the heterotrait triangles of both the monomethod and heteromethod blocks" (Campbell & Fiske, 1959, p. 83). The first requirement provides evidence of convergent validity, while the last three offer evidence of discriminant validity. Table 5 illustrates Campbell and Fiske's (1959) multitrait-multimethod matrix using synthetic data.

Although Campbell and Fiske's (1959) original mititrait-multimethod matrix approach provided "a framework for a rigorous examination of questions concerning convergent and discriminant validity" (Biesanz & West, 2004, p. 850), later studies (Althauser & Heberlein, 1970; Alwin, 1974; Campbell & O'Connell, 1967; Jackson, 1969; Kalleberg & Kluegel, 1975; Krause, 1972) pointed out "three problems" (Schmitt & Stults, 1986, p. 2) with this analysis. Summarizing previous research, Schmitt and Stults (1986) indicated that, first, Campbell and Fiske (1959) did not provide a method for "quantifying the degree to which the criteria were met" (p. 2). Second, Campbell and Fiske's (1959) approach did not allow for "separating method variance from random error" (Schmitt & Stults, 1986, p. 2). Finally, "the Campbell-Fiske criteria were incomplete" (Schmitt & Stults, 1986, p. 2), and the assumptions that there are no correlations between trait and method factors were "implicit in the Campbell-Fiske criteria" (Schmitt & Stults, 1986, p. 2).

Researchers suggested a variety of statistical approaches to address the problems delineated above. As summarized by Schmitt and Stults (1986), these approaches included the analysis of variance (ANOVA) paradigm, which was originally proposed by

Table 5

*A Synthetic Multitrait-Multimethod Matrix*

| | | Method 1 | | | Method 2 | | | Method 3 | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Traits | $A_1$ | $B_1$ | $C_1$ | $A_2$ | $B_2$ | $C_2$ | $A_3$ | $B_3$ | $C_3$ |
| Method 1 | $A_1$ | (.89) | | | | | | | | |
| | $B_1$ | .51 | (.89) | | | | | | | |
| | $C_1$ | .38 | .37 | (.76) | | | | | | |
| Method 2 | $A_2$ | *.57* | .22 | .09 | (.93) | | | | | |
| | $B_2$ | .22 | *.57* | .10 | .68 | (.94) | | | | |
| | $C_2$ | .11 | .11 | *.46* | .59 | .58 | (.84) | | | |
| Method 3 | $A_3$ | *.56* | .22 | .11 | *.67* | .42 | .33 | (.94) | | |
| | $B_3$ | .23 | *.58* | .12 | .43 | *.66* | .34 | .67 | (.92) | |
| | $C_3$ | .11 | .11 | *.45* | .34 | .32 | *.58* | .58 | .60 | (.85) |

*Note.* The validity diagonals are the three sets of italicized values. The reliability diagonals are the three sets of values in parentheses. Each heterotrait-monomethod triangle is enclosed by a solid line. Each heterotrait-heteromethod triangle is enclosed by a broken line. From "Convergent and discriminant validation by the multitrait-multimethod matrix," by D. T. Campbell and D. W. Fiske, 1959, *Psychological Bulletin*, *56*(2), p. 82. Copyright 1959 by the American Psychological Association.

Guilford (1954) and elaborated by Stanley (1961), Boruch and Wolins (1970), and Boruch, Larkin, Wolins, and MacKinney (1970); the generalized proximity function (Hubert & Baker, 1978); a partial correlation method (Schriesheim, 1981a, 1981b); exploratory factor analysis (Golding & Seidman, 1974; Tucker, 1966, 1967); and

confirmatory factor analysis (Bagozzi et al., 1991; Biesanz &West, 2004; Brown, 2006; Cole, 1987; Kenny & Kashy, 1992; Langer, Wood, Bergman, & Piacentini, 2010; Reichardt & Coleman, 1995; Schmitt & Stults, 1986). Developments in confirmatory factor analysis allowed researchers to address problems concerning "*the degree* of convergent and discriminant validity in MTMM matrices in a unified and coherent framework" (Biesanz & West, 2004, p. 850). Therefore, confirmatory factor analysis of multitrait-multimethod matrices was used in the present study to estimate construct validity of M.B.A. rankings.

The present study examined the degree of convergent and discriminant validity of the *Bloomberg Businessweek* full-time M.B.A. rankings using multitrait-multioccasion, confirmatory factor analysis that presents a variation of the traditional multitrait-multimethod matrices approach used in examining validity of the measurement of individual traits. Under this research design, multiple administrations of the same method over time (the *Bloomberg Businessweek* rankings of M.B.A. programs in 2008, 2010, and 2012) served as a method factor. The emphasis on the longitudinal aspect of construct validity in this research is accounted for by the fact that very few studies of rankings investigated "the dynamics and trajectory of their developments" (O'Connell, 2013, p. 719). Although O'Connell's (2013) observation was made with respect to the global rankings of universities, this statement can be applied to describe current research in the sphere of rankings of M.B.A. programs in the United States.

The choice of the *Bloomberg Businessweek* rankings for this type of analysis is

explained by three reasons. First, *Bloomberg Businessweek* is the oldest[15] rankings of business schools in the United States. *Bloomberg Businessweek* magazine has been ranking the top full-time M.B.A. programs in the U.S. since 1988, which gives it more public visibility compared to other ranking systems. Second, the *Bloomberg Businessweek* ranking system (16%) was identified as one of the important factors that influenced the prospective M.B.A. students' decision to apply to a particular school or program (Graduate Management Admission Council, 2012b, p. 27). Only school websites (24%) and *The Financial Times* rankings (17%) were found to be more influential than *Bloomberg Businessweek* in the prospective students' decision to apply to a business school (Graduate Management Admission Council, 2012b, p. 27). Third, the *Bloomberg Businessweek* rankings have three diverse indicators of quality: Student Survey Rank, Employer Survey Rank, and Intellectual Capital Rank, in their methodology, which suits the requirement of the minimum number of three traits in the multitrait-multimethod, confirmatory factor analysis models (Brown, 2006; Widaman, 1985).

Other ranking systems of M.B.A. programs, namely, *The Financial Times*, *Forbes*, and *U.S. News & World Report* are not good datasets for specifying CFA models

---

[15] *Bloomberg Businessweek* has been ranking the top full-time M.B.A. programs in the United States since 1988 and the international programs since 2000. *U.S. News & World Report* started producing rankings of full-time M.B.A. programs in 1990. *Forbes* and *The Financial Times* published their first full-time M.B.A. rankings in 1999, and *The Economist* entered the rankings industry only in 2002.

due to several reasons. First, the *Forbes* ranking system uses four financial indicators to calculate the 5-year M.B.A. gain or the return on investment in an M.B.A. degree. When treated as traits in CFA models, these indicators will not demonstrate discriminant validity as they represent only financial aspects of attaining an M.B.A. degree. Second, *U.S. News & World Report,* which has eight indicators or traits of business schools' quality, and *The Financial Times* with 20 diverse indicators of business schools' quality, present certain challenges from the CFA modeling standpoint. As Tanaka (1987) noted, the appropriateness of sample size should be linked to the size of the model to be estimated, more specifically, "to the number of parameters estimated in the model" (p. 143). Because the sample sizes in these rankings are small, the ratio of sample size to the number of free parameters estimated does not meet the minimum requirement of 5 to 1[16], as recommended by Bentler and Chou (1987). Moreover, the number of free parameters exceeds the sample size for some CFA models, which will distort parameter estimation. Nevertheless, in case the results of CFA of multitrait-multioccasion matrices for the

---

[16] The problem of the appropriate sample size in factor analysis and structural equation modeling has been examined in a variety of studies over many years. Various rules of thumb for the minimum sample size were proposed, depending on different aspects of design. MacCallum, Widaman, Preacher, and Hong (2001), for example, proposed that sample sizes might be smaller than a 4 to 1 ratio of subjects to variables when communalities are high. The researchers (MacCallum et al., 2001) recommended a 20 to 1 subjects-to-variables ratio "when communalities are low and factors are not highly overdetermined" (p. 636).

*Bloomberg Businessweek* rankings of M.B.A. programs were not strong, *The Economist* rankings of M.B.A. programs might be considered for the analysis, although this ranking system is the least influential as only 6% of the prospective M.B.A. students reported that *The Economist* rankings influenced their decision to apply to a business school (Graduate Management Admission Council, 2012b, p. 27).

The multitrait-multimethod, confirmatory factor analysis models used in this study included both *correlated methods* and *correlated uniqueness* models (Brown, 2006; Marsh, 1989; Widaman, 1985). The correlated methods CFA models are built to meet the following specifications: (1) there must be at least three traits (T) and three methods (M) in the model; (2) T x M indicators are used to define T + M latent factors; (3) each indicator is specified to load on its trait factor and its method factor while all other cross-loadings are fixed to zero; (4) correlations among trait factors and among methods factors are freely estimated, but the correlations between trait and method factors are usually fixed to zero; and (5) uniquenesses of indicators, which represent variance in the indicators that is not explained by the trait and method factors, are freely estimated but cannot be correlated with the uniquenesses of other indicators (Brown, 2006; Widaman, 1985). As a result, "each indicator is considered to be a function of trait, method, and unique factors" (Brown, 2006, p. 218).

The correlated uniqueness models' specifications, as outlined by Brown (2006) based on the previous studies (Kenny & Kashy, 1992; Marsh 1989), which are taken into account in this study while building the CFA models, include (1) there must be at least two traits (T) and three methods (M); (2) each indicator is specified to load on one trait

factor while all other cross-loading are fixed to zero; and (3) correlations among trait factors are freely estimated. The correlated uniqueness models are also based on "the assumption of zero correlations among methods" (Brown, 2006, p. 229). In the correlated uniqueness model, "method effects are estimated by specifying correlated uniquenesses (errors) among indicators based on the same assessment method rather than by method factors" (Brown, 2006, p. 221).

When applying multitrait-multimethod, confirmatory factor analysis models to assess construct validity, "factor loadings, variances, and relationships among the latent traits are used to interpret the level or degree of convergent and discriminant validity" (Biesanz & West, 2004, p. 853). "The parameter estimates produced by correlated methods solutions provide seemingly straightforward interpretations with regard to construct validity; for example, large trait factor loadings suggest favorable convergent validity, small or nonsignificant method factor loadings imply an absence of method effects, and modest trait factor intercorrelations suggest favorable discriminant validity" (Brown, 2006, p. 227).

With regard to the parameter estimates produced by correlated uniqueness CFA models, "trait factor loadings that are large and statistically significant would be viewed in support of convergent validity" (Brown, 2006, p. 221). At the same time, "large correlations among the trait factors would be indicative of poor discriminant validity" (Brown, 2006, p. 221). "The presence of appreciable method effects is reflected by correlated uniquenesses among indicators assessed by the same method that are moderate or greater in magnitude" (Brown, 2006, p. 221).

Based on the general assumptions for correlated methods and correlated uniqueness CFA models delineated above, the specifications for CFA models in this study included three traits and three methods. Traits were operationalized as three indicators in the *Bloomberg Businessweek* rankings of full-time M.B.A. programs, namely, Student Survey Rank, Employer Survey Rank, and Intellectual Capital Rank. Methods were operationalized as the *Bloomberg Businessweek* rankings in 2008, 2010, and 2012. Because this study focused on three traits (T) and three methods (M), the CFA models had nine indicators (T x M) and six latent factors (T + M) as presented in Figure 2A. Each indicator was specified to load on its trait factor and its method factor while all other cross-loadings were fixed to zero. Correlations among three trait factors and among three method factors were freely estimated, but the correlations between trait and method factors were fixed to zero. In terms of interpretations of the results of the proposed correlated methods CFA models, large trait factor loadings suggested favorable convergent validity, small or nonsignificant method factor loadings indicated an absence of method effects, and modest trait factor intercorrelations suggested favorable discriminant validity. For correlated uniqueness CFA models, moderate or high correlations of uniquenesses among indicators assessed by the same method (the *Bloomberg Businessweek* rankings) implied the presence of method effects.

When testing the significance of a set of parameters using structural equation modeling, it is also common "to compare the fit of two models, one of which includes the set of parameters and the other of which excludes only the set of parameters to be tested. Comparisons of this sort may be formulated in order to test the degree of convergent and

discriminant validity, as well as the degree of method variance, exhibited by a set of measures" (Widaman, 1985, p. 9).

Table 6 presents a description of the data used to test for construct validity of the *Bloomberg Businessweek* rankings utilizing multitrait-multioccasion, confirmatory factor analysis models. Using observations from business schools that have rankings for all three years, the sample size totals 42 schools.

Table 6

*Data Sources for Multitrait-Multioccasion, Confirmatory Factor Analysis Models*

| Year | *Bloomberg Businessweek* |
|---|---|
| 2008 | 30[17] schools |
| 2010 | 57 schools |
| 2012 | 63 schools |

The type of the data provided in the *Bloomberg Businessweek* rankings presented a certain challenge for multitrait-multioccasion, confirmatory factor analysis. The three indicators, Student Survey Rank, Employer Survey Rank, and Intellectual Capital Rank, that serve as trait factors are rank-ordered data. There is an on-going debate in the current methodological literature about the appropriateness of several correlational

[17] Although *Bloomberg Businessweek* ranked top 30 U.S. full-time M.B.A. programs in 2008, it provided information on indicators used to calculate the rank, Graduate Survey Rank, Corporate Survey Rank, and Intellectual Capital Rank, for 45 schools, including the unranked schools that were in the second tier. Therefore, the sample size totaled 45 schools in 2008.

analysis options for ordinal data. While some researchers (Brown, 2006; Flora & Curran, 2004) recommend that the polychoric correlation coefficient be used for examining relationships among ordinal variables in confirmatory factor analysis, others (Choi, Peters, & Mueller, 2010; Ekström, 2011) argue that the Spearman rank order correlation coefficient can also be appropriate, especially when the sample size is very small (Choi et al., 2010). The third group of researchers (e.g., Luo, 2011) proposes "a non-parametric polychoric correlation coefficient based on the discrete version of Spearman's rho" (Luo, 2011, p. 20). In light of this controversy, Ekström (2011) demonstrated that "Spearman's rank correlation coefficient can be expressed as a deterministic transformation of the empirical polychoric correlation coefficient, and vice versa" (p. 13), and "the two measures of association are asymptotically equivalent in a certain sense" (p. 13-14). At the same time, Ekström (2011) argued that the empirical polychoric correlation is "more conservative" (p. 14) if ordinal variables are "the result of some form of discretization, such as grouping of values into categories" (p. 14). Based on the fact that the ordinal data in the *Bloomberg Businessweek* ranking system are rank-ordered data and are not aggregated into categories and the sample size ($n = 42$) is small, this study utilized the Spearman rank order correlation coefficient to fit models for CFA of multitrait-multimethod matrices.

Based on Marsh's (1989) taxonomy of structural models for MTMM data adapted from Widaman (1985), this study used six multitrait-multioccasion, confirmatory factor analysis models to assess the degree of convergent and discriminant validity of the *Bloomberg Businessweek* rankings.

1. **The correlated traits, correlated methods model (CTCM)** is presented in Figure 2A. In this model, three indicators of quality of M.B.A. programs in the *Bloomberg Businessweek* rankings serve as traits. Methods are examined from the longitudinal perspective, and the *Bloomberg Businessweek* rankings in 2008, 2010, and 2012 serve as method factors. Three traits are correlated as well as three methods. Nevertheless, the correlations between trait and method factors are fixed to zero. Uniquenesses of indicators are not correlated. The correlated traits, correlated methods model (CTCM) is the least restricted CFA model to which other more restricted models are compared. As Widaman (1985) proposed, by contrasting the fit of this model with "properly specified, more restricted models, tests of convergent and discriminant validity and of the amount of method-related variance are possible" (p. 9).
2. No traits, correlated methods model (NTCM) is presented in Figure 2B. In this model, there are no traits, and method factors are represented by the Bloomberg Businessweek rankings in 2008, 2010, and 2012, which are correlated. Uniquenesses of indicators are not correlated. This model is nested under the CTCM model. Comparison of this model, where only correlated method factors are present, to the CTCM model provides a test of convergent validity. Comparison of the CTCM and the NTCM allows researchers "to obtain an estimate and test of the covariance among measures

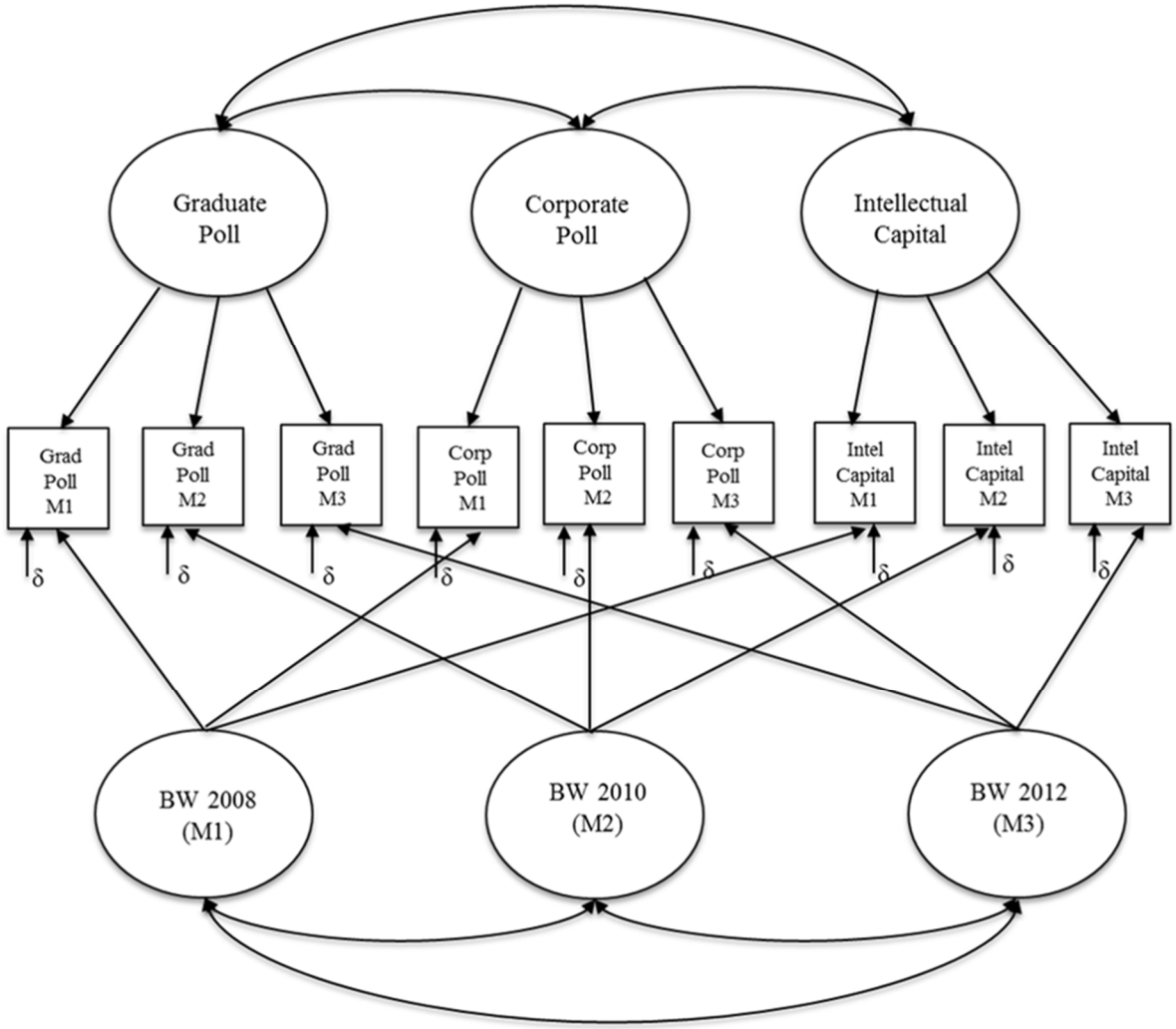


*Figure 2A*. The correlated traits, correlated methods model (CTCM).

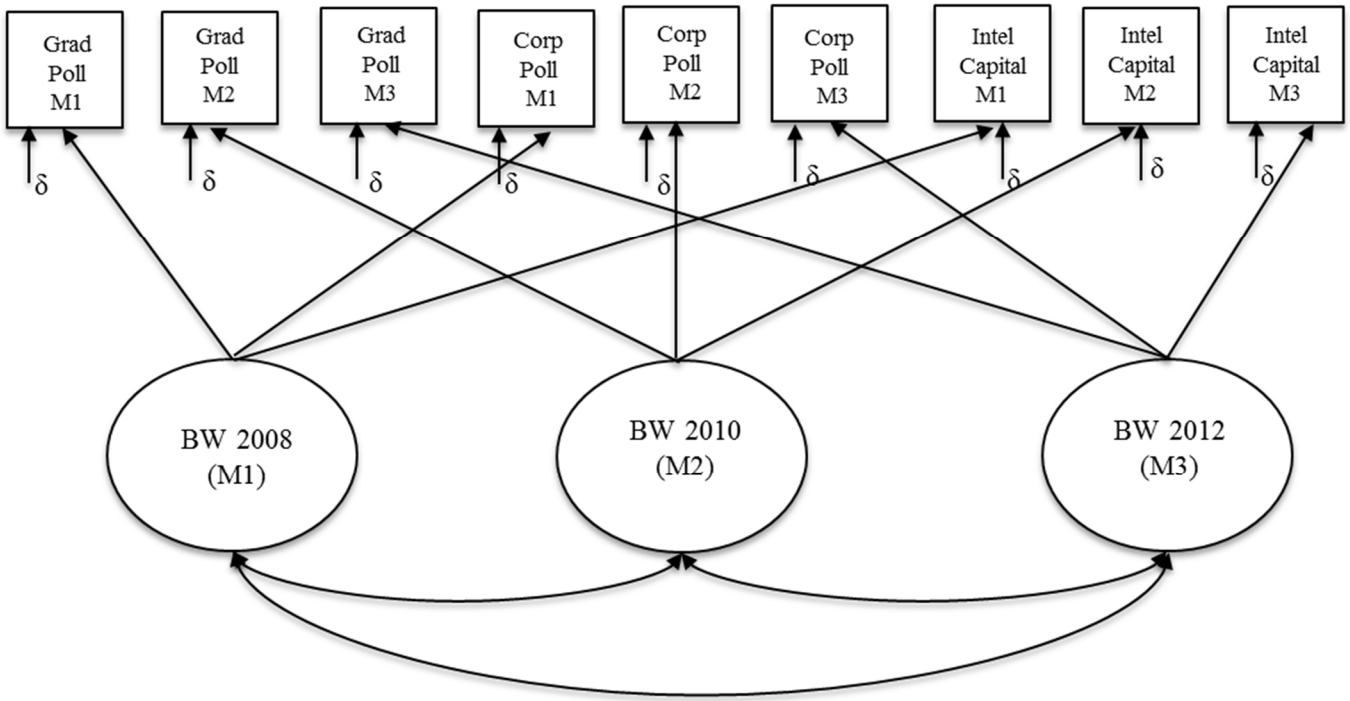


*Figure 2B.* No traits, correlated methods model (NTCM).

uniquely explained by trait factors, which represent convergent validity" (Widaman, 1985, p. 9).

3. **The orthogonal traits, correlated methods model (OTCM)** is presented in Figure 2C. This model is nested under the CTCM model. Comparison of this model, where trait factors are orthogonal, to the CTCM model yields a test of discriminant validity of the traits. If the CTCM fits the data the same as the OTCM does, there is little evidence of discriminant validity among the trait constructs, and the traits do not diverge meaningfully.
4. **The correlated traits, orthogonal methods model (CTOM)** is presented in Figure 2D. This model is nested under the CTCM model. Comparison of this model, where method factors are orthogonal, to the CTCM model offers a test of "the amount of method covariance among the methods" (Widaman, 1985, p. 10), which is a test of discriminant validity of the methods. If model CTCM provides better fit than does model CTOM, then "the method factors display significant levels of covariation" (Widaman, 1985, p. 10). The CTOM is also nested under the CTCU. "The comparison between the CTOM and the CTCU presents a test of whether the method effects are unidimensional" (Biesanz & West, 2004, p. 855), which means that "the method effects associated with each method form a single latent method factor" (Marsh & Grayson, 1995, p. 186).

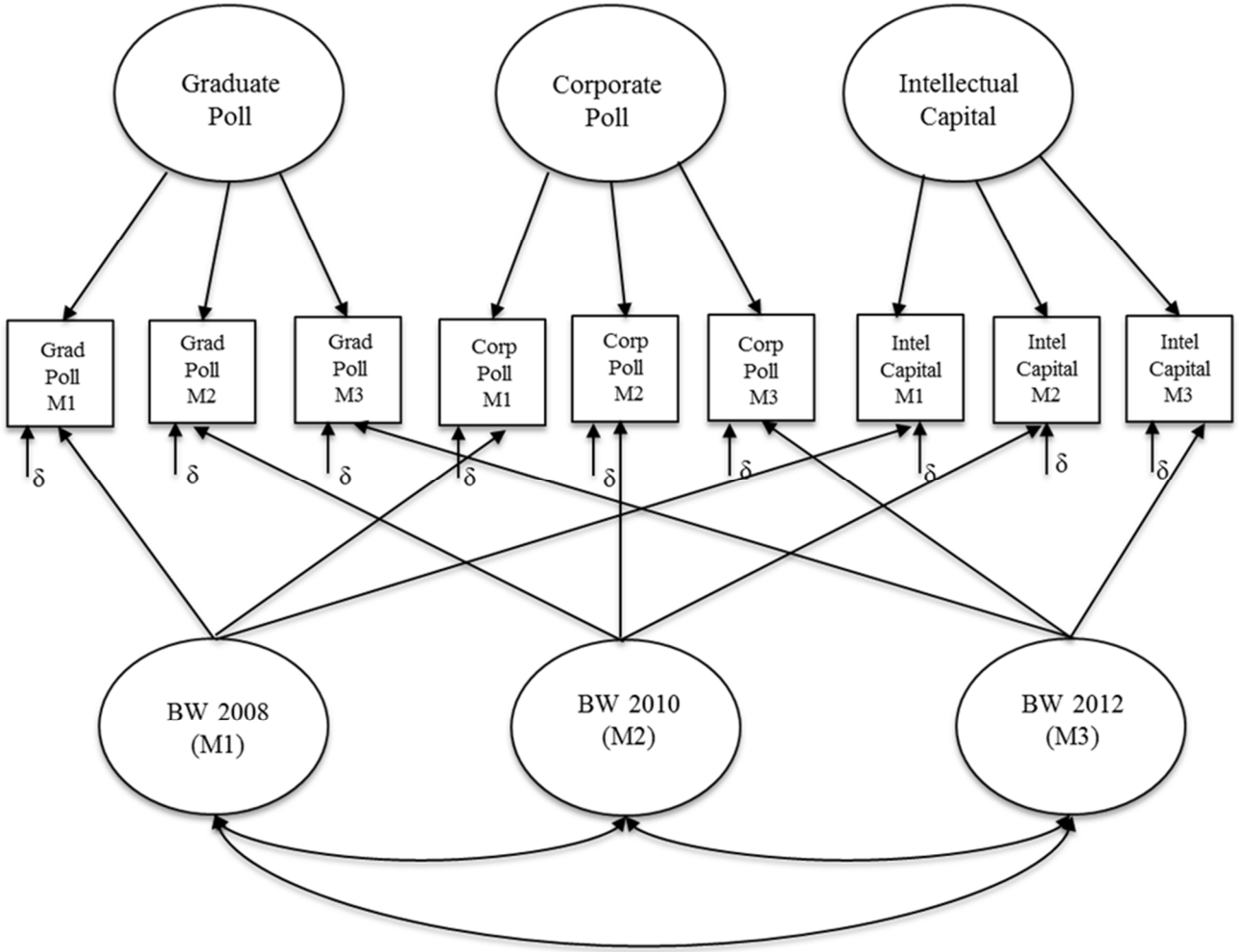


*Figure 2C.* The orthogonal traits, correlated methods model (OTCM).

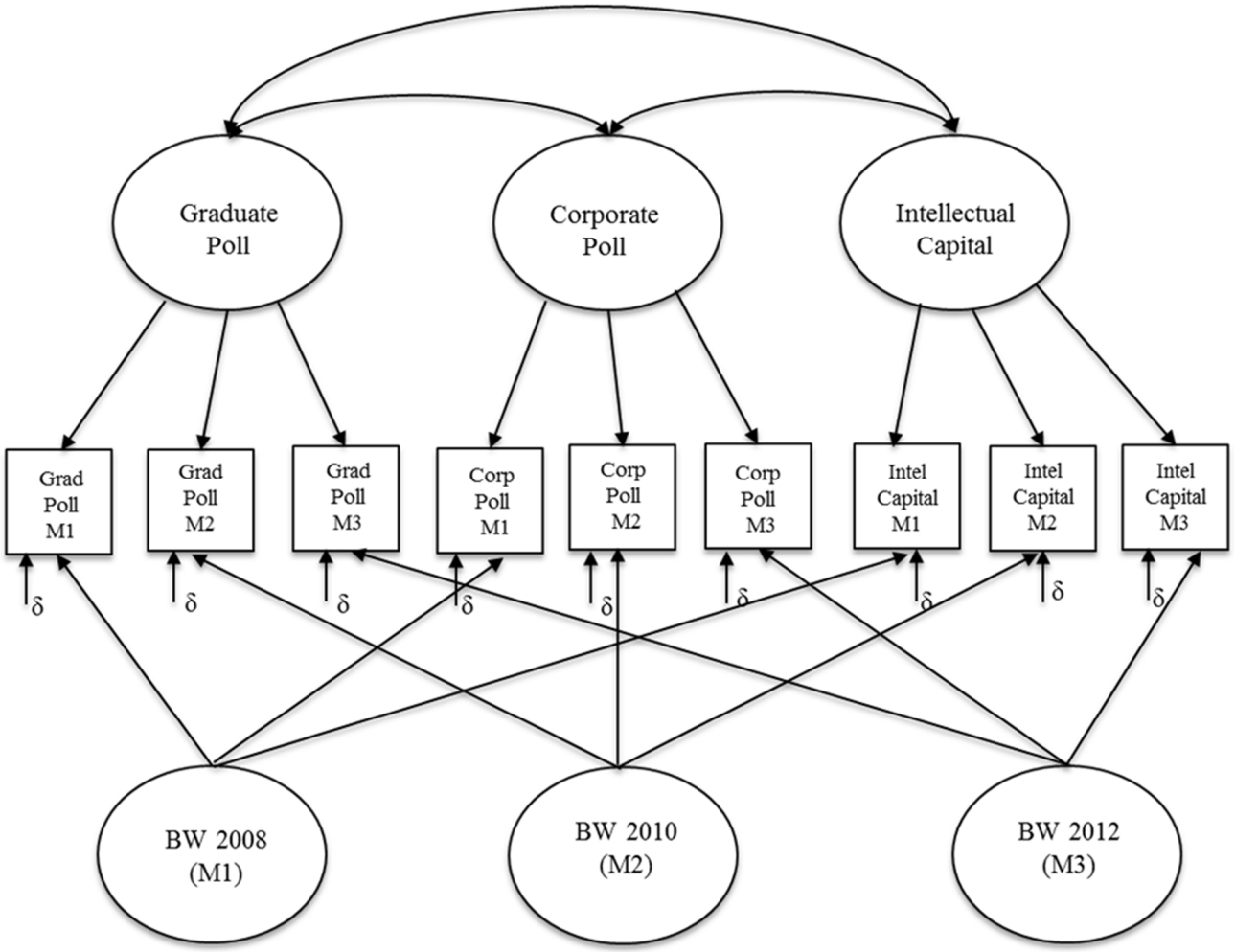


*Figure 2D.* The correlated traits, orthogonal methods model (CTOM).

5. **The correlated traits, correlated uniqueness model (CTCU)** is presented in Figure 2E. The CTCU model assumes that methods effects are uncorrelated with each other (Brown, 2006).
6. **The orthogonal traits, correlated uniqueness model (OTCU)** is presented in Figure 2F. The OTCU model assumes that method effects are uncorrelated with each other. This model is nested under the CTCU model. Comparison of the OTCU and the CTCU presents a test of orthogonality of the trait factors or discriminant validity of the traits.

LISREL 9.10 (Jöreskog & Sörbom, 2014) statistical software was used for multitrait-multioccasion, confirmatory factor analysis in this research.

**Research Question 3**

**Conundrums of reputation: Anchoring effects in the full-time M.B.A. rankings in the United States.** This part of the research addresses the third research question and examines whether the measure of quality of full-time M.B.A. programs is independent of the measure of quality of the higher education institutions in which the M.B.A. programs are located. In other words, this research question delves into the problem of a potential anchoring (or halo) effect of rankings of higher education institutions and their reputation on rankings of their business schools and their reputational assessments.

According to the tenets of anchoring theory, "people make estimates by starting from an initial value that is adjusted to yield the final answer" (Tversky & Kahneman, 1974, p. 1128). Tversky and Kahneman (1974) further claimed that "adjustments are

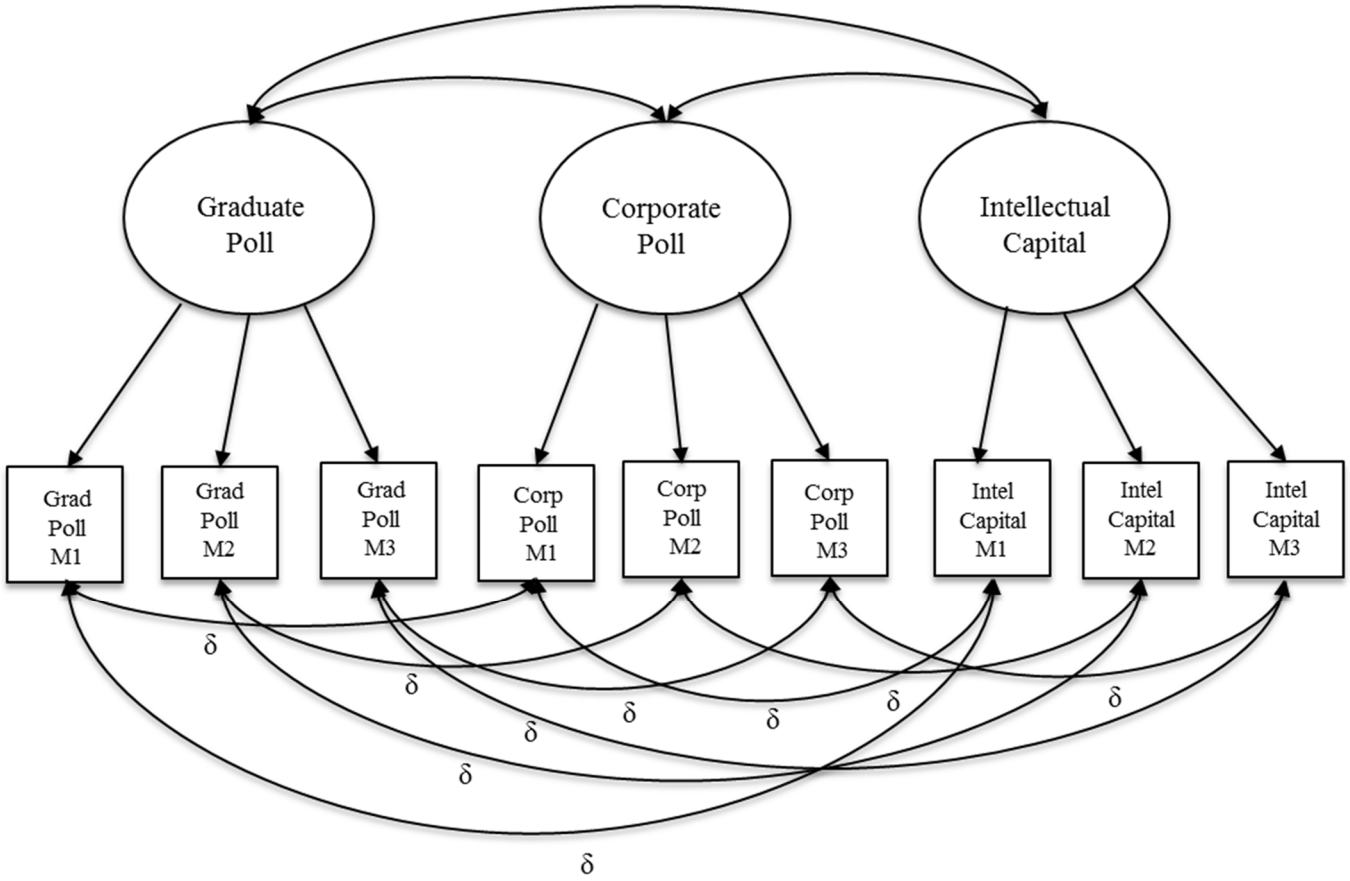


*Figure 2E.* The correlated traits, correlated uniqueness model (CTCU).

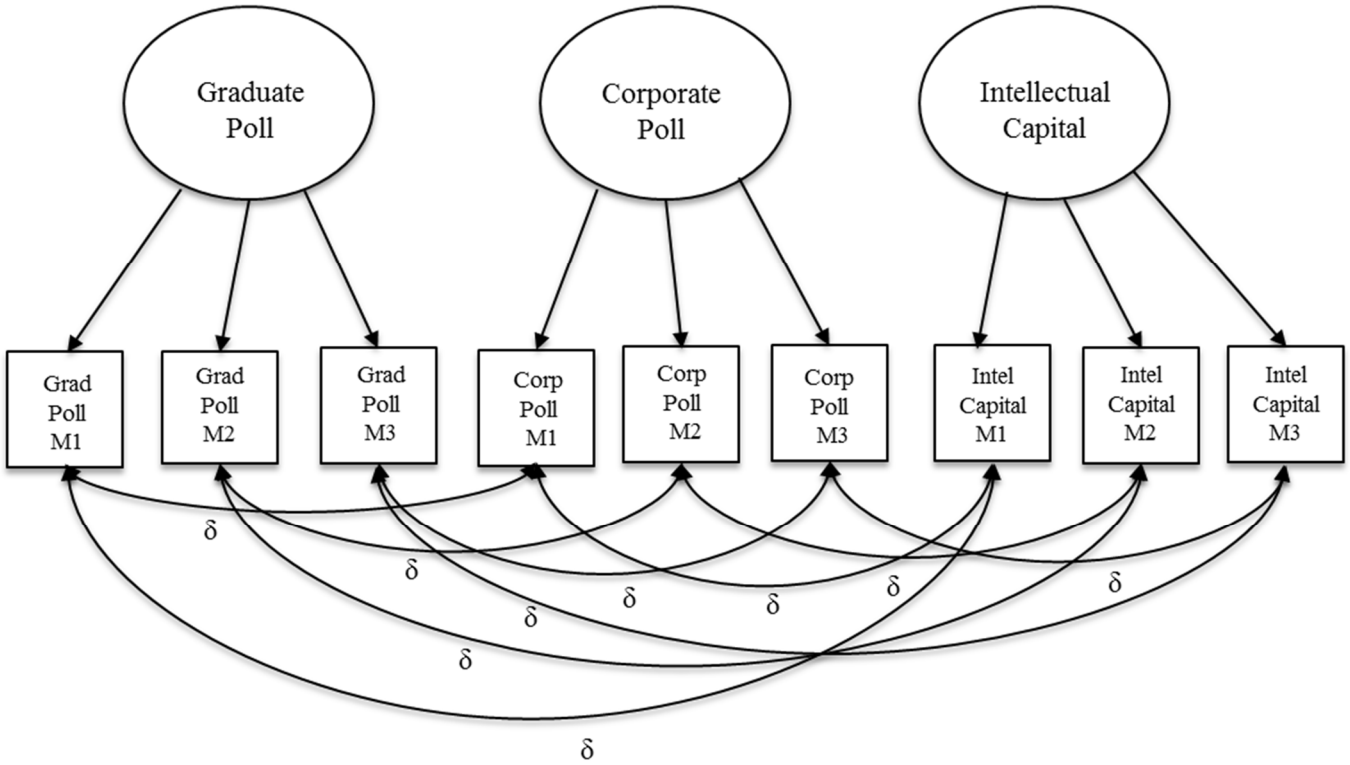


*Figure 2F.* The orthogonal traits, correlated uniqueness model (OTCU).

typically insufficient" (p. 1128), and "different starting points yield different estimates, which are biased toward the initial values" (p. 1128). As a result, the initially presented value exerts a differential influence on people's judgment or decision-making.

Although there has been "nearly 40 years worth of research on the anchoring effect which has proved to be extremely robust" (Furnham & Boo, 2011, p. 41) in psychology, survey design, economics, law, and other social studies, the phenomenon of the anchoring effect has entered the research paradigm in higher education quite recently. A few recent reputational studies provided indirect evidence of the anchoring effect in university and graduate and professional school rankings. Bastedo and Bowman (2010), for example, found the powerful influence of the introductory year of the *U.S. News & World Report* undergraduate rankings on future college reputational assessments, which was "independent of changes in organizational quality and performance and even of prior peer assessments of reputation" (p. 163). A more recent study by Bowman and Bastedo (2011), which focused on the anchoring effects of the *Times Higher Education Supplement* world university rankings, corroborated their previous findings. According to Bowman and Bastedo (2011), "the initial *THES* rankings influenced peer assessments of reputation in subsequent surveys, but second-year rankings were not related to changes in reputation in the third year" (p. 431).

In the field of professional education, similar findings were demonstrated in the studies of business, law, and medical school rankings. Consistent with the anchoring perspective, Stake (2006) utilized a 15-year data set (1990 - 2004) and showed that the *U.S. News & World Report* law school rankings predicted changes in schools' reputation

during the earlier years, and the differences between ranking and reputation became smaller over time. Recent research on business school reputation suggests more complex models of reputation than those found in Stake's (2006) study. Rindova et al. (2005), for example, viewed reputation as a construct that consisted of perceived quality and prominence and reflected recruiters' opinion. Safón (2009) hypothesized business school reputation as "a unidimensional construct" (p. 204) that incorporated perceptions of various stakeholders, such as business schools' deans, M.B.A. candidates, and recruiters of M.B.A. graduates. Recent research on business school reputation also focuses on a wider scope of ranking systems as predictors of business school reputation, such as *Bloomberg Businessweek* (Rindova et al., 2005) and *Bloomberg Businessweek, Forbes, The Financial Times*, *The Wall Street Journal*, and *U.S. News & World Report* (Safón, 2009). Although different in their methodological approaches and data sample, Rindova et al. (2005) utilized data from *Businessweek's* 1998 rankings, and Safón (2009) used data from the above mentioned five M.B.A. rankings in 2003, both studies (Rindova et al., 2005; Safón, 2009) concurred that media rankings significantly predicted business school reputation.

A few studies lent an interesting perspective to the examination of the anchoring effect in higher education by focusing on the relationship between the reputation of professional schools and the reputation of universities with which they are affiliated. In a study of the reputation of 94 American medical schools, Cole and Lipton (1977), for example, reported that while "reputation stems from functionally appropriate performance" (p. 662), such as research and publications, pre-eminence of faculty,

training and research grants, size of full-time faculty, and effectiveness of training, "there is some evidence of a ceiling effect (Harvard) and a halo effect for schools affiliated with universities having national reputations" (p. 662).

A recent study by Safón (2012) explored antecedents of business school reputation from "two competing perspectives: the meritocratic and the institutional" (p. 169). Under meritocratic perspective, Safón (2012) theorized reputation as a construct that combines institutional attributes and performance measures, including selectivity, size, student/faculty ratio, program value, media ranking, and research performance. The institutional perspective, according to Safón (2012), posited "the importance of the institutional context in explaining reputation" (p. 169), such as business school past reputation and parent university reputation. Using data from four ranking systems, *Bloomberg Businessweek, U.S. News & World Report, Quacquarelli Symonds (QS) World University Rankings*, and *University of Texas at Dallas Top 100 Business School Research Rankings*, Safon's (2012) study revealed that past reputations of both a business school and its parent university had almost the same significant impact on their current reputations, with standardized beta coefficients of .90 ($p < .001$) and .91 ($p < .001$), respectively. Moreover, Safón (2012) demonstrated that "the reputations of schools and their parent universities are influenced reciprocally" (p. 177). The impact of the past parent university reputation on the current business school reputation and vice versa was approximately the same, with standardized beta coefficients of .07 ($p < .01$) and .10 ($p < .05$), respectively. Safón's (2012) finding of the reciprocal influence of the reputations of business schools and their parent universities is of special interest as it opens new

avenues in organizational research in higher education, highlighting the problem of interdependence of different structural parts at universities as "loosely coupled systems" Weick (1976, p. 1).

Stemming from previous research on university and graduate and professional school reputation, briefly reviewed above, this study used a different approach to examining the anchoring effects in the full-time M.B.A. rankings in the United States. While previous research primarily drew on the earlier reputational assessments or rankings as "the starting value to inform their judgment" (Bowman & Bastedo, 2011, p. 433) about subsequent reputational assessments or rankings, this study explored the anchoring effect of both rankings and reputation of a higher education institution on rankings and reputation of its structural part, a full-time M.B.A. program.

Out of the five full-time M.B.A. ranking systems in the U.S.A. that are under analysis in this research, only *U.S. News & World Report* meets the data criteria for this research question. First, *U.S. News & World Report* provides both National University Rankings and Best Graduate Schools: Schools of Business rankings. Second, both National University Rankings and School of Business rankings provided by *U.S. News & World Report* have a reputational component in their research methodologies. The other four ranking systems either do not provide university rankings or do not have a reputational component in their methodologies. Taking into account these data limitations, this study examined whether and under what conditions the rank and reputation of a university, as determined by the *U.S. News & World Report* National University Rankings, affect the rank and reputation of a full-time M.B.A. program at this

university, as determined by the *U.S. News & World Report* Schools of Business rankings.

The research hypothesis follows from the preceding discussion. Because the *U.S. News & World Report* National University Rankings serve as an anchor for peer raters' and recruiter raters' judgment in the rankings of M.B.A. programs, the overall university rankings will be positively associated with both future M.B.A. program rankings and their reputational assessments, controlling for earlier reputational assessment components in the National University Rankings.

All data were obtained from the *U.S. News & World Report* National University Rankings and the *U.S. News & World Report* Best Graduate Schools: Schools of Business rankings. The recent three years of the rankings were used (i.e., 2011, 2012, and 2013)[18].

The conceptual framework that drove the variable selection and analysis for this study was adopted from Bowman and Bastedo (2011) and modified in accordance with the research hypothesis in this study. In order to examine how a rank and reputation of a university, as measured by the *U.S. News & World Report* National University Rankings, affect a rank and reputation of a full-time M.B.A. program, as measured by the *U.S. News & World Report* Best Graduate Schools: Schools of Business rankings, structural equation modeling (SEM) (Hoyle, 1995; Maruyama, 1998) was used. Figures 3A, 3B, and 3C present the path analysis models from the conceptual framework guiding this

[18] The data for this study were collected in 2013.

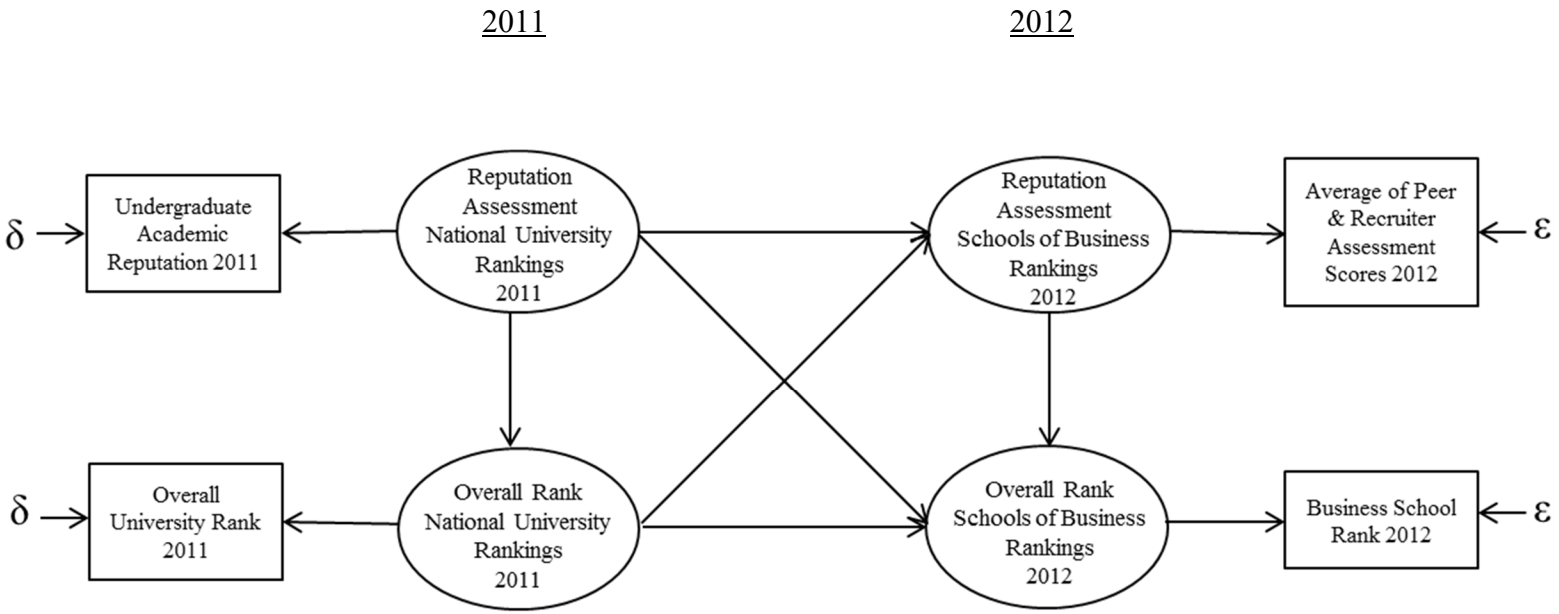


*Figure 3A.* Conceptual framework: Structural Equation Model 1.

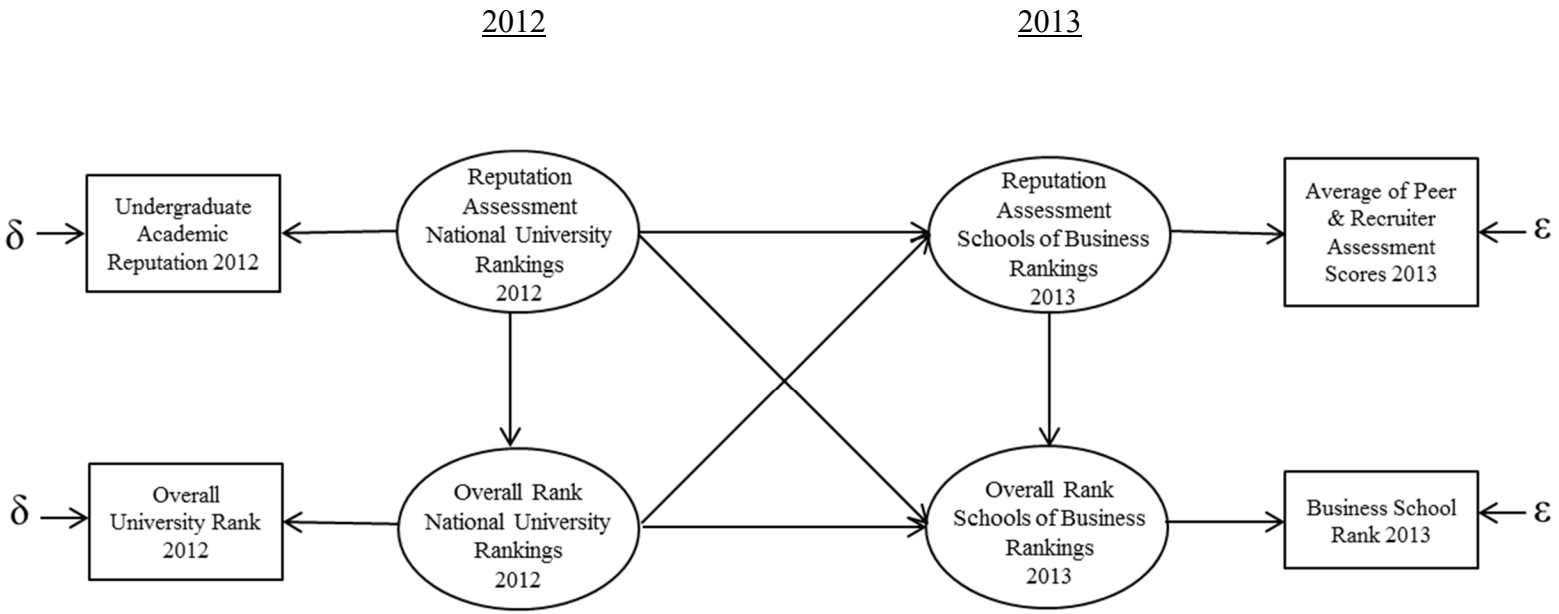


*Figure 3B.* Conceptual framework: Structural Equation Model 2.

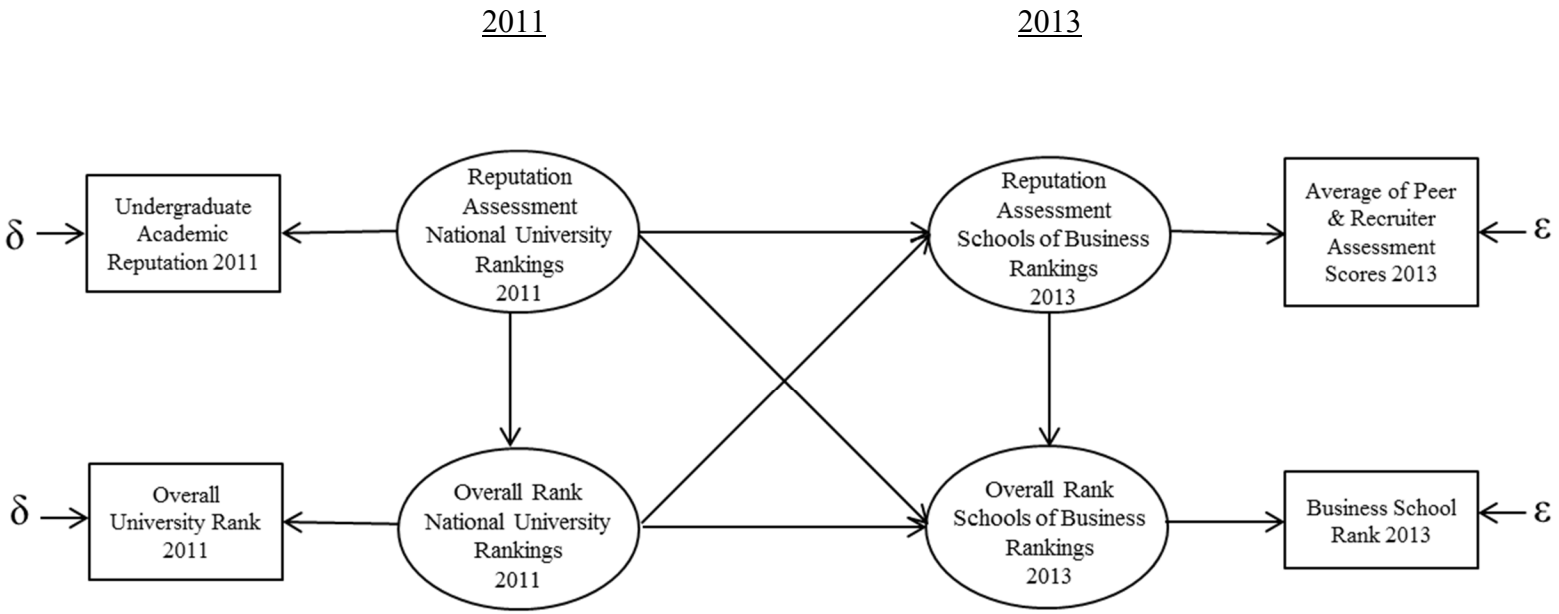


*Figure 3C.* Conceptual framework: Structural Equation Model 3.

study. In Figures 3A and 3B, the models test whether the overall university rankings and their reputation will be positively associated with the M.B.A. program rankings and their reputational assessments in the consecutive year (lag 1). In Figure 3C, the model tests for the effect of the overall university rankings and their reputation in 2011 on the M.B.A. program rankings and their reputational assessments in 2013, which represents a longer lag (lag 2). The model in Figure 3C seeks to test whether the effect of the overall university rankings and their reputation on the M.B.A. program rankings and their reputational assessments increases or decreases over time.

The endogenous variables in the models are the overall rank of universities in the National University Rankings at Time 1, the overall rank of business schools at Time 2, and the reputational assessment component of business school rankings at Time 2. The overall rank of business schools is represented by the rank of business schools in a particular year. The reputational assessment of business schools is operationalized as the average of the peer assessment score and the recruiter assessment score in a particular year. The exogenous variable in the models is the reputational assessment component in the National University Rankings at Time 1. Both the reputational assessment of national universities and the reputational assessment of business schools are quality indicators used to calculate the overall ranks of a university and a business school, representing 22.5% and 40% of a final overall score, respectively. Therefore, the path analysis models in Figures 3A, 3B, and 3C include a causal link between the reputational assessment and the rank. Table 7 presents a description of the data used to test for a potential anchoring

effect of rankings of higher education institutions and their reputation on rankings of their business schools and their reputational assessments.

Table 7

*Data Sources for Structural Equation Models*

| Year | *U.S. News & World Report Best Graduate Schools: Schools of Business Rankings* | *U.S. News & World Report National University Rankings* |
|---|---|---|
| 2011 | 72 schools | 72 universities |
| 2012 | 69 schools | 69 universities |
| 2013 | 72 schools | 72 universities |

Using observations from business schools and their parent universities that have rankings for all three years, the sample size totals 64 schools. All variables used in the analysis are continuous. Undergraduate academic reputation index in the university rankings has the maximum score of 100. Peer assessment and recruiter assessment scores in the business school rankings have the maximum score of 5. Reputational assessment scores of business schools in the models are calculated as an average of the peer and recruiter assessment scores. Overall scores are used to represent the corresponding ranks of universities and business schools, with 100 being the maximum score. Although the variances among all variables were very different due to the differences in their measurement scales, the variables were not standardized because LISREL 9.10 (Jöreskog & Sörbom, 2014) statistical software, which was used for path analysis in this research, provides both unstandardized and standardized solutions.

**Research Question 4**

**Are we making money for the entire institution? The impact of rankings of full-time M.B.A. programs on funding streams to institutions of higher education.** This part of the research addresses the fourth research question and explores how rankings of full-time M.B.A. programs are associated with funding streams to institutions of higher education.

Despite the widespread popularity of college rankings and a wealth of research on the effects that rankings have on applications, enrollment decisions, and admission outcomes (Bowman & Bastedo, 2009; McDonough, Antonio, Walpole, & Perez, 1998; Meredith, 2004; Monks & Ehrenberg, 1999; Sauder, & Lancaster, 2006), the impact of rankings on institutions' financial resources is less explored and yields controversial results in empirical studies. Monks and Ehrenberg (1999), for example, found that a one-rank improvement in the *U.S. News & World Report* rankings allowed selective private institutions to raise net tuition by 0.3% the next year. At the same time, while analyzing data for both public and private universities, Meredith (2004) found no evidence that rankings have an impact on the amount of gifts that a university received, which included corporate support, research grants, and alumni donations.

A recent study by Bastedo and Bowman (2011) provided an examination of the impact of the 1998 *U.S. News & World Report* National University Rankings on multiple financial indicators in 2000, 2002, and 2006, including overall research and development funding, federal research and development funding, state/local research and development funding, industry research and development funding, instate tuition and fees, out-of-state

tuition and fees, proportion of alumni donating, total alumni giving, and foundation funding. Bastedo and Bowman's (2011) findings demonstrated that "controlling for previous financial indicators, peer assessment, institutional control, and changes in objective quality, being ranked below Tier 1 adversely affects R&D funding (overall, federal, and industry), the proportion of alumni donating, and out-of-state tuition and fees" (p. 16) in 2006. With respect to financial indicators in 2000 and 2002, significant effects were observed only for overall, federal, and state/local R&D funding, and the effect sizes were smaller than for financial indicators in 2006.

In the field of professional education, Cheslock and Gianneschi (2008) found that an institution's business school rankings, as well as medical school research rankings, might be a strong predictor of giving revenues for the entire institution. As Cheslock and Gianneschi (2008) contended, "Falling one slot in the *U.S. News* rankings for medical research is associated with a drop in current fund gifts per student of close to $20, and institutions outside of the top 50 receive $1,200 to $1,600 less per student than those institutions at the top of the rankings. Similar results occur for MBA rankings in the 2004 regression, but an institution's law school rankings consistently is not a strong predictor of giving revenues" (p. 221).

Based on the promising findings of the impact of university rankings on various revenue streams to a higher education institution in Bastedo and Bowman's (2011) study and the impact of business school rankings on an institution's giving revenues in Cheslock and Gianneschi's (2008) research, this study hypothesized that business school rankings bolster financial revenues for the entire institution. More specifically, the

research hypothesis states that after controlling for prior institutional financial resources, a rank of a business school will have a significant effect on future resource flows from the federal, state and local government, business, institutional funds, nonprofit organizations, and other sources to a research university for research and development.

In order to examine how a rank of a business school impacts funding streams to a higher education institution, structural equation modeling (SEM) (Hoyle, 1995; Maruyama, 1998) was used. The overall rank of a business school was operationalized as a full-time M.B.A. program rank measured by the *U.S. News & World Report* rankings. Figures 4A, 4B, 4C, 4D, 4E, and 4F present the path analysis models from the conceptual framework guiding this study. The endogenous variables in the six models were R&D funding coming from the federal government, R&D funding from the state and local government, R&D funding from institution funds, R&D funding from business, R&D funding from nonprofit organizations, and R&D funding from all other sources in 2011. The exogenous variables in the models were the overall rank of a business school in 2010 and the overall R&D expenditure of a higher education institution in 2010.

The business school data used to test for the impact of rankings of full-time M.B.A. programs on funding streams to institutions of higher education were obtained from the *U.S. News & World Report* Best Graduate Schools: Schools of Business rankings. The 2010 rankings were used, with 51 schools ranked.

Research and development funding (R&D) data were collected from the National Science Foundation's Higher Education Research and Development Survey (HERD) 2011 (National Science Foundation/National Center for Science and Engineering

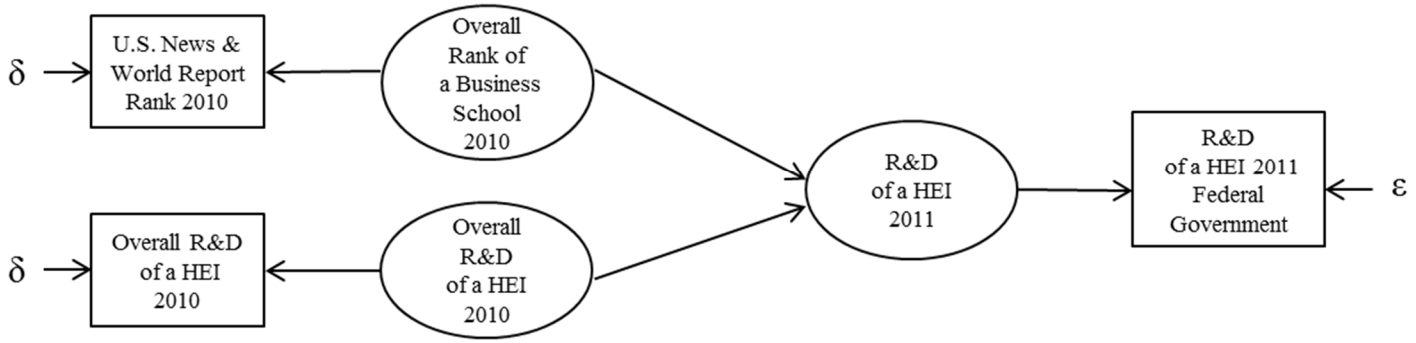


*Figure 4A.* Conceptual framework: Structural Equation Model 1.

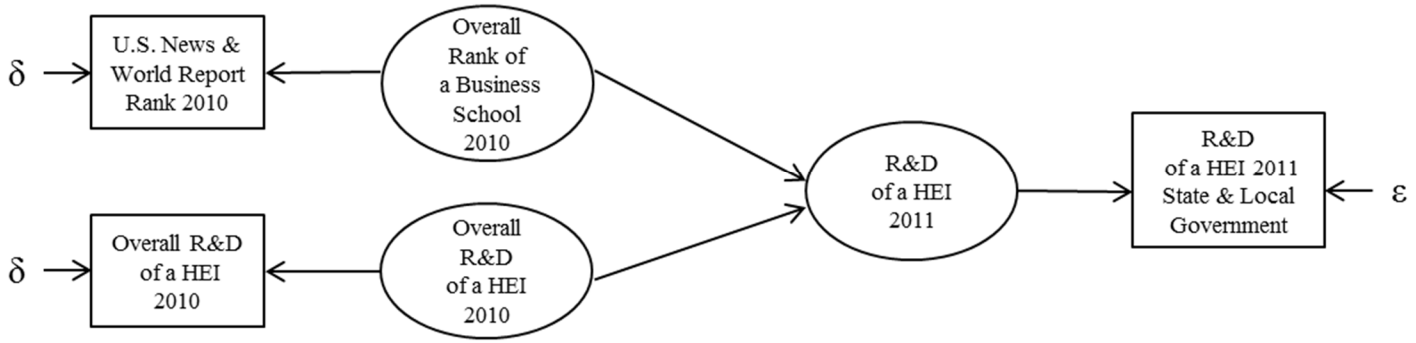


*Figure 4B.* Conceptual framework: Structural Equation Model 2.

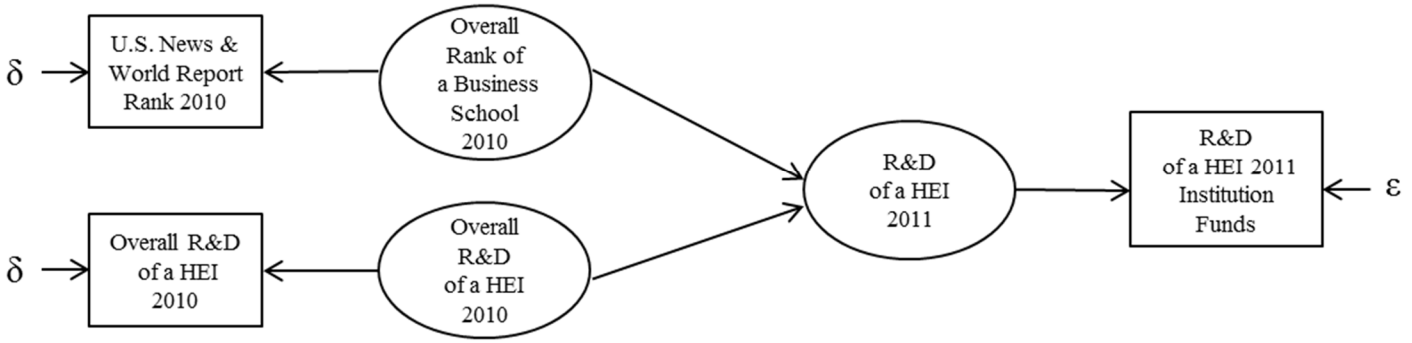


*Figure 4C.* Conceptual framework: Structural Equation Model 3.

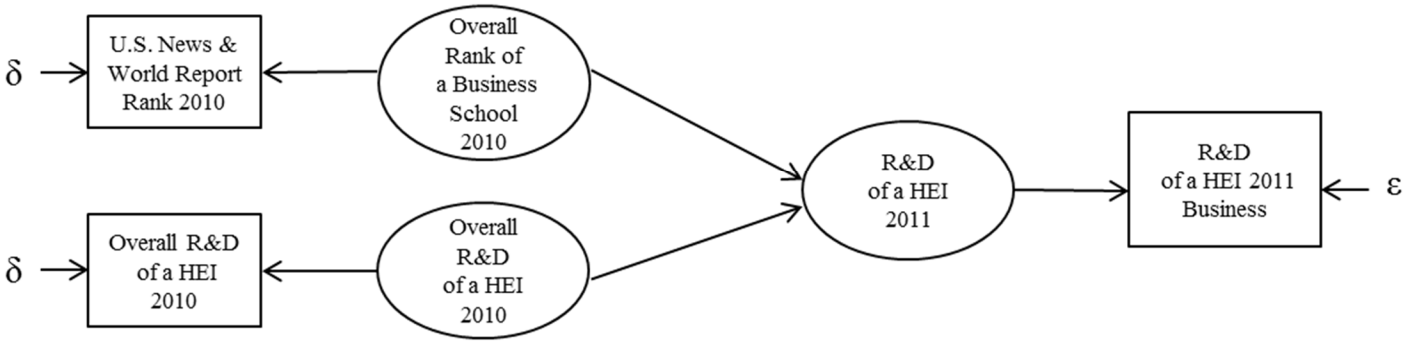


*Figure 4D.* Conceptual framework: Structural Equation Model 4.

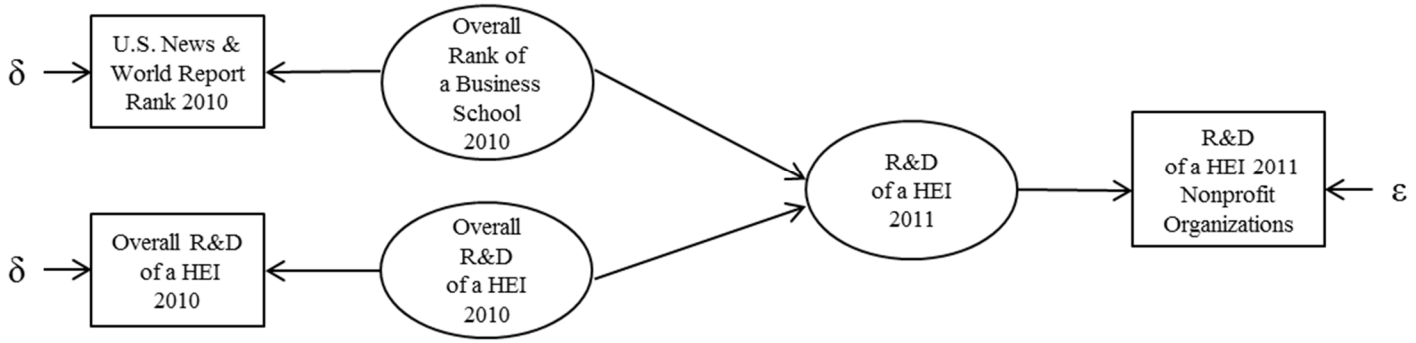


*Figure 4E.* Conceptual framework: Structural Equation Model 5.

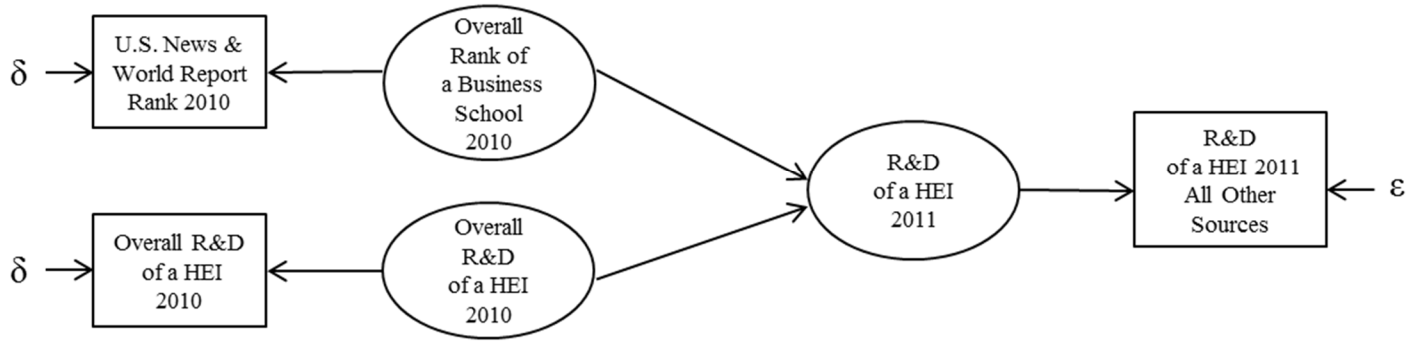


*Figure 4F.* Conceptual framework: Structural Equation Model 6.

Statistics, 2013). The 2011 Higher Education R&D (HERD) Survey (NSF/NCSES, 2013), which is a successor to the Survey of Research and Development Expenditures at Universities and Colleges, gathered information from 912 research-performing academic institutions. The HERD survey was revised and expanded to incorporate R&D within nonscience and engineering (non-S&E) fields in the totals reported throughout the survey. The scope of R&D was also widened to include expenditures on clinical trials and research training grants.

All variables used in the analysis were continuous. Overall scores were used to represent the corresponding ranks of business schools, with 100 being the maximum score. Although the variances among all variables were very different due to the differences in their measurement scales, the variables were not standardized because LISREL 9.10 (Jöreskog & Sörbom, 2014) statistical software, which was used for path analysis in this research, provides both unstandardized and standardized solutions.

# CHAPTER IV: RESULTS

This chapter unveils the results of the study. Based on the research questions, this chapter falls into four sections. Although the statistical procedures utilized to explore each of the four research questions are different, each section in this chapter follows the same pattern. First, each section presents the results of the overall model specification, such as assumptions of multivariate and univariate normality, multicollinearity, and model fit, or other assumptions required by the statistical procedure (the minimum acceptable level of consensus, etc.). Second, each section of this chapter presents the results for each research question per se.

## Research Question 1

**Ranking of rankings: Assessment of the full-time M.B.A. program rankings through the lens of the Berlin Principles on Ranking of Higher Education Institutions.** After the three raters submitted their scores assessing the quality of the five ranking systems of M.B.A. programs in the United States based on the IREG criteria, there was a follow-up discussion to resolve discrepancies in scores and reach the minimum acceptable level of consensus. As shown in Table 8, Krippendorff's α met the minimum acceptable level of the agreement estimate of interrater reliability of .667 (Krippendorff, 2004a) for the *Bloomberg Businessweek*, *The Economist*, *The Financial Times*, and the *U.S. News & World Report* rankings. Krippendorff's α for the *Forbes* ranking system equaled .666. Although this agreement estimate of interrater reliability was lower than the acceptable level, the difference of .001 was very small to invalidate the raters' scores.

Table 8

*Estimates of Interrater Reliability*

| Ranking system | Krippendorff's α |
|---|---|
| *Bloomberg Businessweek* | .675 |
| *The Economist* | .674 |
| *The Financial Times* | .681 |
| *Forbes* | .666 |
| *U.S. News & World Report* | .670 |

I averaged the scores provided by the three raters for each IREG criterion and used weights specified in Table 1 to calculate the total scores for each ranking system. Table 9 provides the total scores for each ranking system that present the assessment of rankings quality by higher education professionals as well as prospective students' evaluation of the importance of various M.B.A. rankings for their decision to enroll in a particular program as reported in the 2012 mba.com Prospective Students Survey (see Figure 1) (Graduate Management Admission Council, 2012b). As demonstrated in Table 9, the total scores for all ranking systems, which were based on the raters' scores, were considerably lower than the maximum total score of 180 specified in the IREG methodology.

There was a moderate positive correlation ($r_s(5) = .62$, $p = .269$) between the rankings of M.B.A. ranking systems calculated based on the prospective students' responses to the survey and the rankings computed based on the scores provided by professionals in academia, but the correlation was not statistically significant.

Table 9

*Comparative Results of Ranking of M.B.A. Program Rankings by Prospective Students and Higher Education Professionals*

| The 2012 mba.com Prospective Students Survey | | | Higher education professionals' assessment based on the IREG methodology | | |
|---|---|---|---|---|---|
| Rank | % of students | Ranking system | Rank | Score | Ranking system |
| 1 | 17% | *The Financial Times* | 1 | 130.98 | *The Financial Times* |
| 2 | 16% | *U.S. News & World Report* | 2 | 114.33 | *Bloomberg Businessweek* |
| 2 | 16% | *Bloomberg Businessweek* | 3 | 111.67 | *The Economist* |
| 4 | 8% | *Forbes* | 4 | 92.68 | *U.S. News & World Report* |
| 5 | 6% | *The Economist* | 5 | 92.31 | *Forbes* |

Table 9 shows that both groups demonstrated agreement about the value of M.B.A. rankings only with regard to two ranking systems – *The Financial Times* and *Bloomberg Businessweek*. As indicated in Table 9, *The Financial Times* ranking system topped the list, followed by the *Bloomberg Businessweek* ranking system, in both rankings. The results of the rankings of M.B.A. ranking systems demonstrated differences in prospective students' and higher education professionals' opinions about M.B.A. rankings, which may indicate that the two groups have different understanding of the quality of business education. This potential misunderstanding of the priorities and goals

between the two groups may become an obstacle to the improvement of business education.

**Research Question 2**

**Reliability of the full-time M.B.A. rankings in the United States: Estimating predictability of changes of the rankings.** As shown in Table 10, all the four ranking systems exhibited a highly significant positive first-order autocorrelation of the error terms. These results indicate that ranking systems of the full-time M.B.A. programs are predictable, and the future rankings will demonstrate stability of either an upward or downward trend. Nevertheless, the four ranking systems differed in the magnitude of their first-order autocorrelations, which implies that some rankings are more predictable than others. The *Forbes* ranking system had the highest first-order autocorrelation in both balanced $\rho$[19] = .80, $p < .001$ and unbalanced $\rho = .76$, $p < .001$ data sets. First-order autocorrelations for *U.S. News & World Report* appeared to be very close in their magnitude, with $\rho = .77$, $p < .001$ in balanced and $\rho = .73$, $p < .001$ in unbalanced data sets, to those of the *Forbes* ranking system. *The Financial Times* ranking system had a somewhat lower first-order autocorrelation than *U.S. News & World Report* in the balanced data set $\rho = .70$, $p < .001$ and a notably lower first-order autocorrelation $\rho = .57$,

---

[19] Statistical textbooks use different symbols to define first-order autocorrelation. Kutner, Nachtshein, Neter, and Li (2005), for example, use $\rho$ symbol in the generalized multiple regression model when the random terms follow a first-order autoregressive process (pp. 484-485), whereas Ljung and Box (1978) use $r$ in their formula of the Ljung-Box test statistics (p. 298).

Table 10

*First-Order Autocorrelations for the Four Ranking Systems of M.B.A. Programs*

| Ranking system | Number of observations | First-order autocorrelation | Ljung-Box[20] statistic | *p* |
|---|---|---|---|---|
| *U.S. News & World Report*: | | | | |
| balanced data | 546 | .77 | 326.97 | $p < .001$ |
| unbalanced data | 768 | .73 | 407.82 | $p < .001$ |
| *Bloomberg Businessweek*: | | | | |
| balanced data | 175 | .19 | 6.35 | $p = .012$ |
| unbalanced data | 270 | .20 | 11.41 | $p = .001$ |
| *The Financial Times*: | | | | |
| balanced data | 392 | .70 | 193.64 | $p < .001$ |
| unbalanced data | 765 | .57 | 245.12 | $p < .001$ |
| *Forbes*: | | | | |
| balanced data | 264 | .80 | 171.94 | $p < .001$ |
| unbalanced data | 385 | .76 | 226.00 | $p < .001$ |

$p < .001$ in the unbalanced data set. The *Bloomberg Businessweek* ranking system exhibited the lowest first-order autocorrelations in both balanced and unbalanced data sets, with $\rho = .19$, $p = .012$ and $\rho = .20$, $p = .001$, respectively.

[20] The Ljung and Box (1978) modified test weighs the squared autocorrelation at lag $k$ by $(n + 2)/(n - k)$ $(k = 1, ..., m)$. The Ljung-Box test statistic is computed as

$$Q(\hat{r}) = n(n+2)\sum_{k=1}^{m}(n-k)^{-1}\hat{r}_k^2$$

(Ljung & Box, 1978, p. 298).

Figures 5 and 6 show autocorrelation functions of the error terms at all lags in both balanced and unbalanced data sets of the four full-time M.B.A. programs ranking systems. From Figures 5 and 6, it is evident that the *Forbes* and the *U.S. News & World*

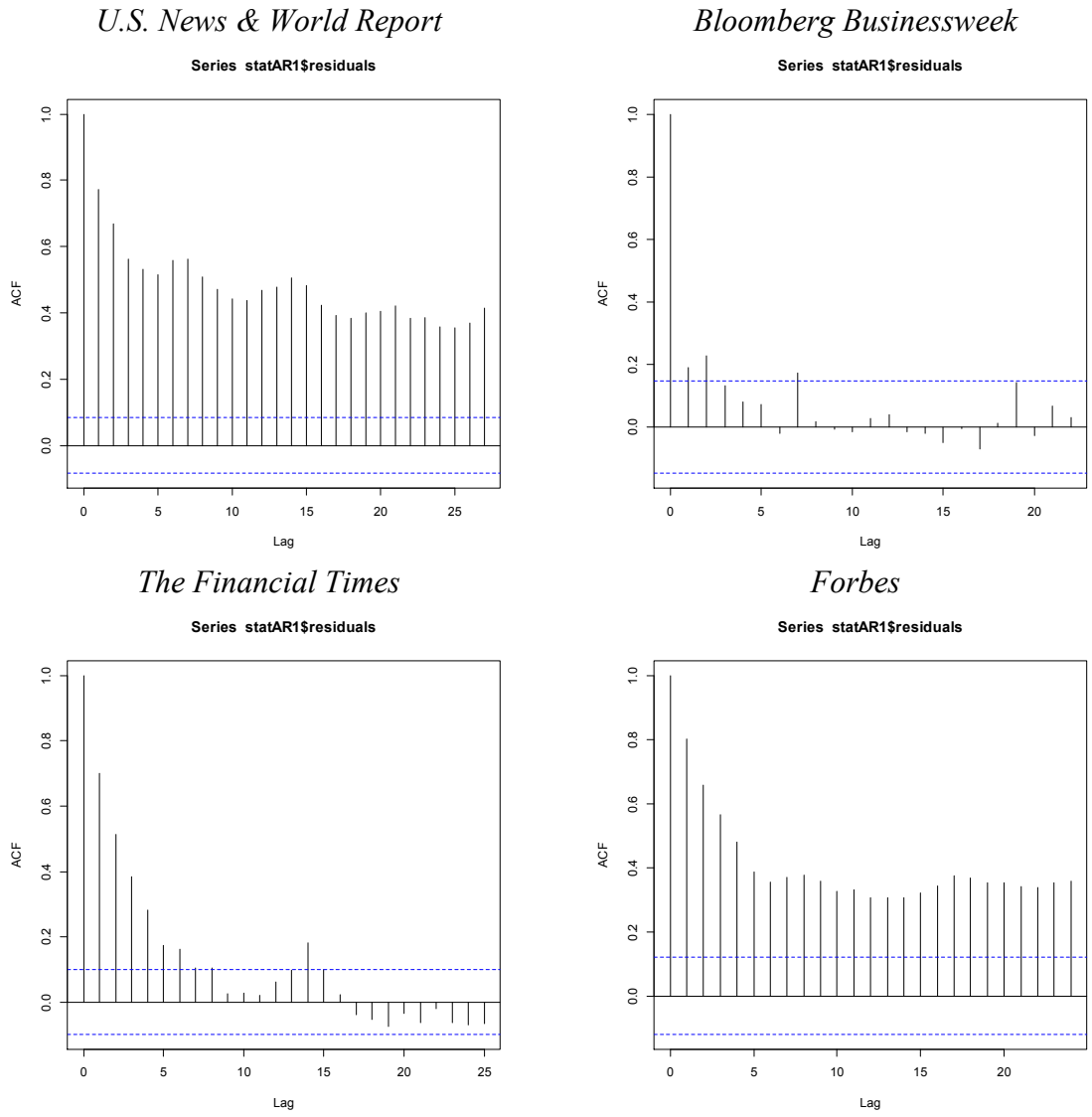


*Figure 5.* Autocorrelation function of error terms for the four full-time M.B.A. rankings with balanced data.

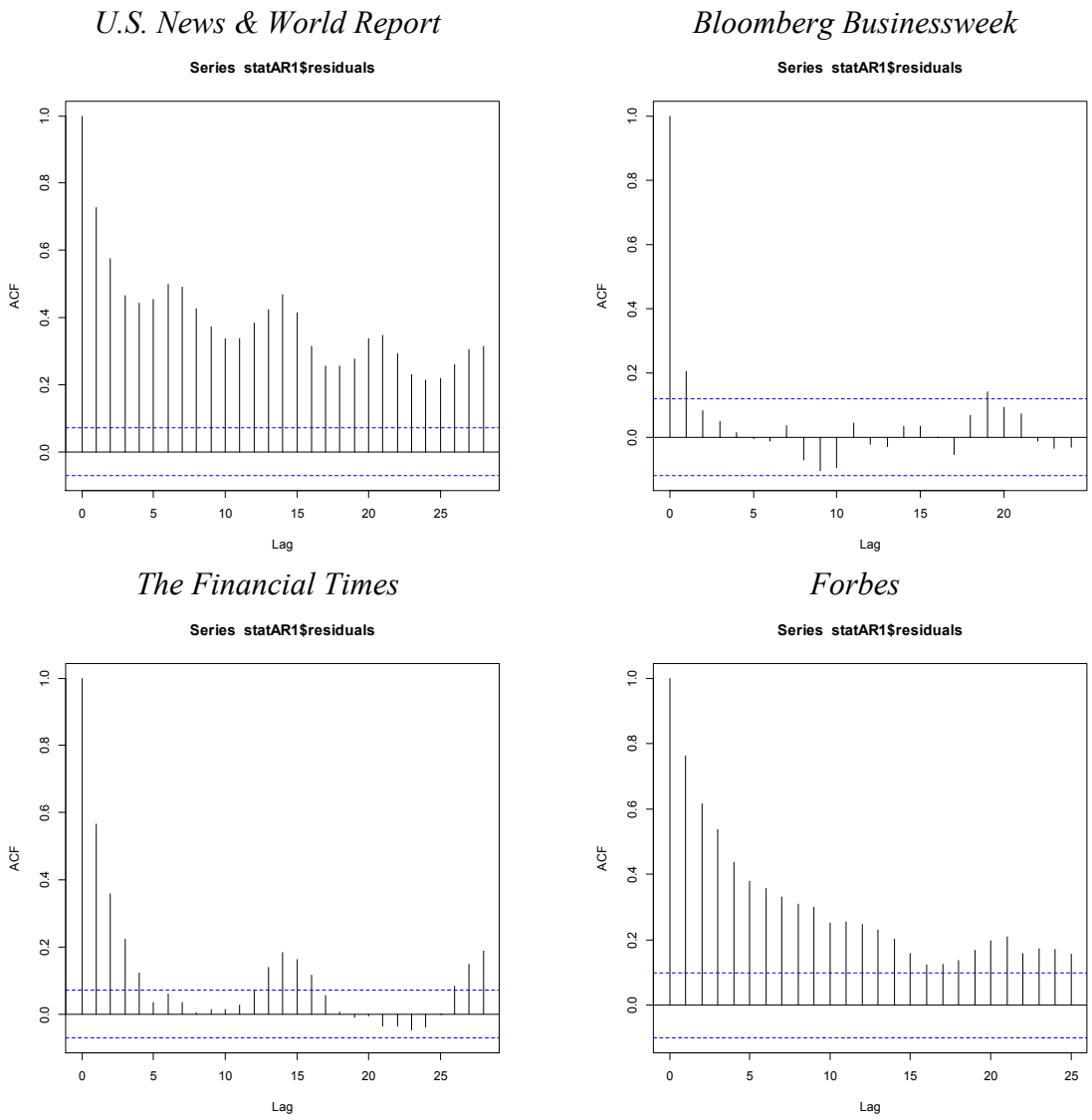


*Figure 6.* Autocorrelation function of error terms for the four full-time M.B.A. rankings with unbalanced data.

*Report* rankings exhibited only positive statistically significant autocorrelations of the error terms at all lags. *Bloomberg Businessweek* and *The Financial Times* rankings

demonstrated both positive and negative autocorrelations of error terms at different lags, but only positive autocorrelations were statistically significant[21].

Figures 7 and 8 present the data plots[22] for the four ranking systems of full-time M.B.A. programs for balanced and unbalanced data sets, respectively. The visual inspection of the plots yields the same conclusions as the statistical results. As shown in the data plots, a portion of business schools in each of the four ranking systems exhibited a more or less noticeable trend over time. The magnitude of the trend and the number of business schools demonstrating the trend varied from rankings to rankings. The visual inspection of the data plots in Figure 7 demonstrates that the *Forbes* rankings had the most business schools that showed a trend over time, followed by the *U.S. News & World Report* and *The Financial Times* rankings. In other words, errors associated with a given time period carried over into the ensuing time period and created a trend in rankings, which was demonstrated by the significant positive first-order autocorrelations in the four

---

[21] In the plots of autocorrelations, "The dashed lines correspond to $-2/\sqrt{n}$ and $+2/\sqrt{n}$, since autocorrelations are declared to be statistically significantly different from zero if they are less than $-2/\sqrt{n}$ or greater than $+2/\sqrt{n}$ (i.e., if they are more than two standard errors away from zero)" (Sheather, 2009, p. 308).

[22] Ranks of M.B.A. programs (Rank) are plotted over time (Year), 2000-2013. The line segments connect the ranks of each M.B.A. program over time. The data plot for the *U.S. News & World Report* rankings in Figure 7, for example, demonstrates that the rank of the top 20 M.B.A. programs exhibited very little variation over time. At the same time, lower ranked programs showed large variations in rankings.

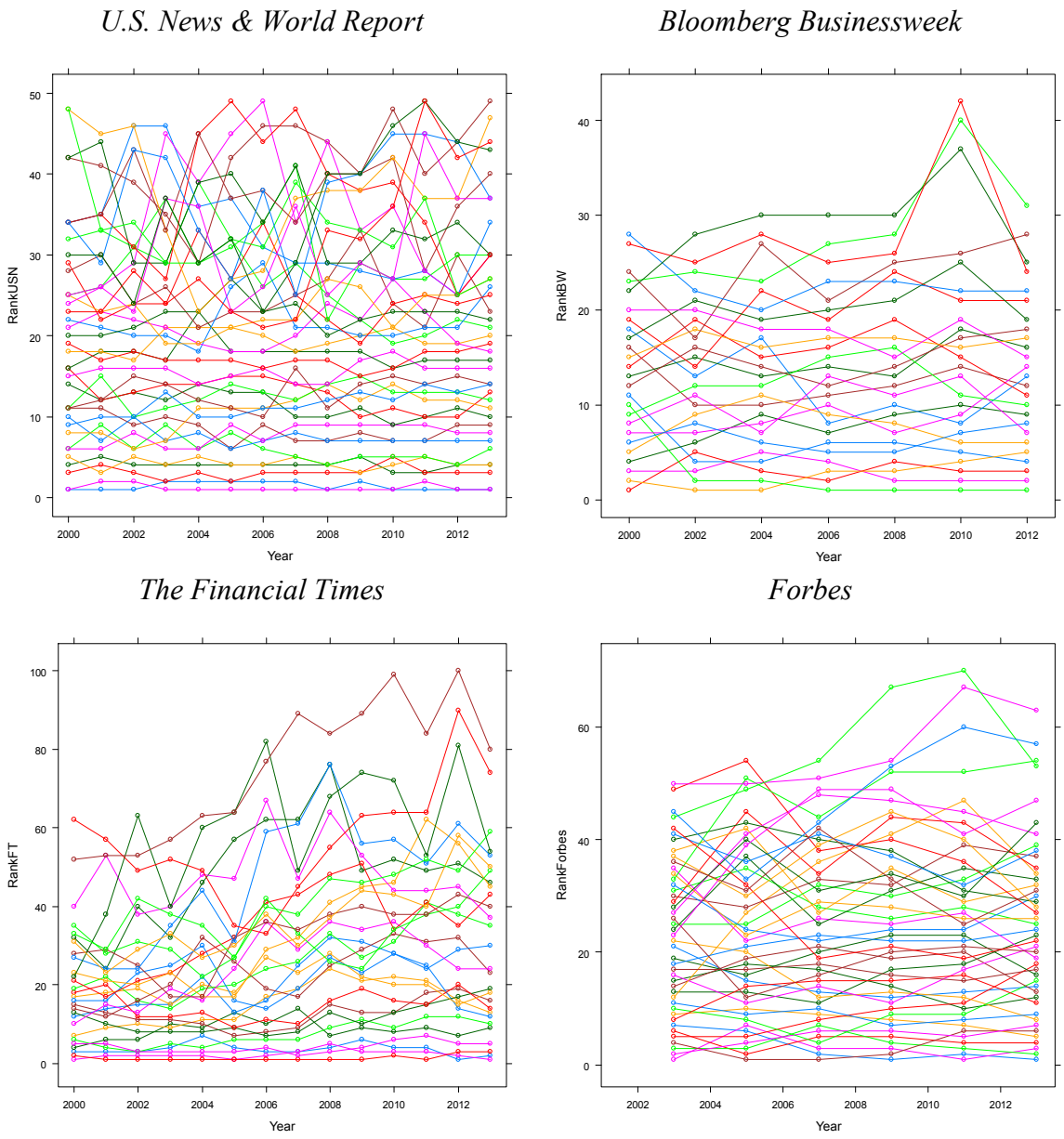


*Figure 7.* Data plots for the four full-time M.B.A. rankings with balanced data, by ranking year.

ranking systems. Data plots in Figure 8 demonstrate the same order of ranking systems in terms of the number (from highest to lowest) of business schools that exhibited a trend over time: *Forbes*, *U.S. News & World Report*, *The Financial Times*, and *Bloomberg Businessweek*.

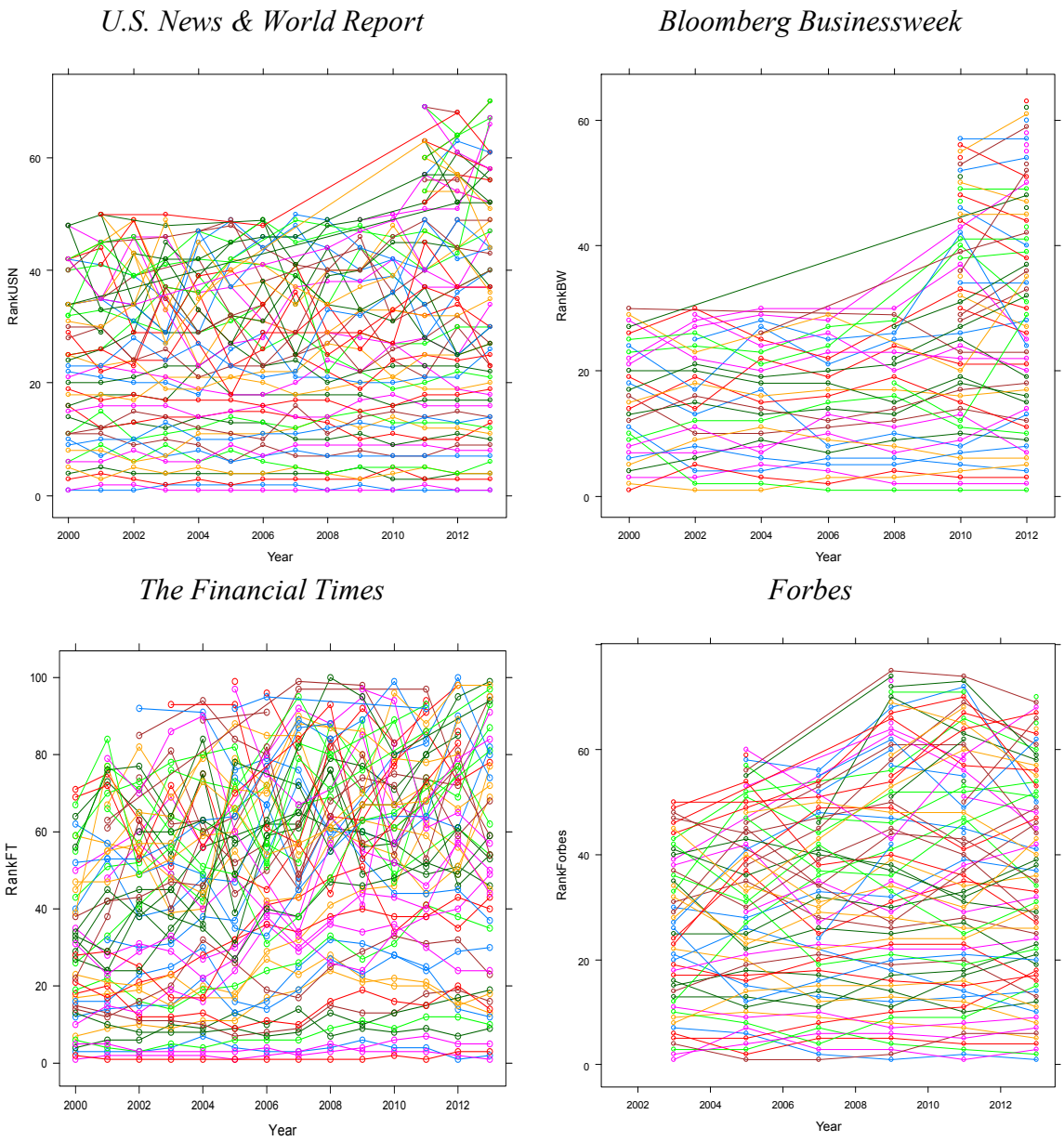


*Figure 8.* Data plots for the four full-time M.B.A. rankings with unbalanced data, by ranking year.

Although the first-order autocorrelations of error terms were statistically significant for all the rankings under analysis in both balanced and unbalanced data sets, and there is no statistical procedure to understand the difference in magnitude of these

first-order autocorrelations, there are some possible explanations of the statistical results. Researchers (Asteriou & Hall, 2011; Greene, 2012; Gujarati, 1995) identified a number of causes of serially correlated errors that include omitted variables, misspecification of the model (such as ignoring nonlinearity), measurement errors, unnecessary manipulation of data (averaging series, interpolation, extrapolation, etc.), data transformation, inertia, cobweb phenomenon, etc. While the variables in the models used to calculate the rank in the four ranking systems are specified by their media publishers, and there is no possibility to test what variables are missing in order to improve the model fit, it is logical to assume that the *Forbes* ranking formula, which consists of only financial variables, exhibits the most lack of fit. As a result, the narrow definition of business school quality in the *Forbes* rankings, which is operationalized by only financial variables, probably explains the highest first-order autocorrelation. Possible misspecification of the models in terms of the incorrect functional form, which is a limitation in this study, might also account for the differences in the magnitude of the first-order autocorrelations. In order to ensure consistency, the autoregressive (AR1) covariance model was chosen to fit to the data for the four ranking systems. Nevertheless, the nature of the time series trend might be different for *Bloomberg Businessweek*, *Forbes*, *The Financial Times*, and *U.S. New & World Report* and cause variations in the magnitude of the first-order autocorrelations as a result of ignoring nonlinearity in the models. Finally, measurement error might account for the differences in the magnitude of the first-order autocorrelations for the four rankings. Measurement error in rank calculation can occur for two major reasons. First, measurement error can

occur in the process of data collection resulting from the quality of the survey instruments used to poll business school administrators, alumni, and recruiters, from sampling procedures, nonresponse, bias, etc., and the quality of the data provided by business schools. Second, measurement error may result from complex data manipulation in the process of calculating the overall rank. As media rankings are not accountable to the public for the quality of their rankings product, it is difficult to gauge the amount of measurement error in the four ranking systems. However, it is possible to speculate that the *Bloomberg Businessweek* rankings have the smallest percent of measurement error based on the fact that this ranking system exhibits the smallest first-order autocorrelation of the error terms.

**Reliability of the full-time M.B.A. rankings in the United States: Estimating reliability as stability of the ranking scores.** Table 11 presents the Spearman rank order correlations between business school rankings published in the seven issues of *Bloomberg Businessweek* between 2000 and 2012. The correlations were computed using pairwise deletion for the sample sizes of M.B.A. programs, ranging from 26 to 67 programs. While pairwise deletion has disadvantages, it was chosen for the analysis because of two reasons. First, pairwise deletion allows calculating the intercorrelation matrix using all available data. Second, different data sets of M.B.A. programs for different years reflect the dynamics of rankings evolution that gradually became more democratic and changed in focus from a small number of top selective M.B.A. programs (30 programs in 2000) to include 57 programs in the 2010 and 63 programs in the 2012 *Businessweek* rankings.

As demonstrated in Table 11, the lag 1 correlations[23] ranged in magnitude from .87 to .97, with a weighted average correlation of .93[24]. Therefore, on average, 86% of the variance in the current *Businessweek* year ranking was shared with the previous year's ranking.

Table 11

*Spearman Rank Order Correlations Between Businessweek Rankings Across Years (2000 – 2012)*

| Variables | 2000 | 2002 | 2004 | 2006 | 2008 | 2010 | 2012 |
|---|---|---|---|---|---|---|---|
| 2000 | 1.00 | | | | | | |
| 2002 | .87 | 1.00 | | | | | |
| 2004 | .89 | .92 | 1.00 | | | | |
| 2006 | .88 | .94 | .94 | 1.00 | | | |
| 2008 | .86 | .91 | .94 | .97 | 1.00 | | |
| 2010 | .84 | .92 | .93 | .94 | .94 | 1.00 | |
| 2012 | .87 | .89 | .93 | .92 | .90 | .93 | 1.00 |

*Note.* All values are statistically significant, $p < .001$.

[23] The sample sizes for correlations between ranking years 2000-2002, 2002-2004, 2004-2006, 2006-2008, 2008-2010, and 2010-2012 equaled 29, 29, 29, 26, 30, and 53, respectively.

[24] The weighted average Spearman rank order correlation was computed using Fisher's *z* transformation, which is recommended "when averaging correlation coefficients, particularly when sample size is small" (Silver & Dunlap, 1987, p. 146).

Table 12 presents the Spearman rank order correlations between business school rankings published in the six issues of *The Economist* between 2007 and 2013, excluding 2009 due to unavailability of data. The correlations were computed using pairwise deletion for the sample sizes of M.B.A. programs, ranging from 38 to 55 programs. As shown in Table 12, the lag 1 correlations[25] varied in magnitude from .95 to .97, with a weighted average correlation of .96. Thus, on average, 92% of the variance in the current *Economist* ranking was shared with the previous year's ranking.

Table 12

*Spearman Rank Order Correlations Between The Economist Rankings Across Years (2007 – 2013)*

| Variables | 2007 | 2008 | 2010 | 2011 | 2012 | 2013 |
|---|---|---|---|---|---|---|
| 2007 | 1.00 | | | | | |
| 2008 | .97 | 1.00 | | | | |
| 2010 | .91 | .95 | 1.00 | | | |
| 2011 | .89 | .91 | .97 | 1.00 | | |
| 2012 | .90 | .92 | .93 | .95 | 1.00 | |
| 2013 | .91 | .90 | .91 | .92 | .97 | 1.00 |

*Note.* All values are statistically significant, $p < .001$.

Table 13 shows the Spearman rank order correlations between business school rankings published in the 14 issues of *The Financial Times* between 2000 and 2013. The correlations were calculated using pairwise deletion for the sample sizes of M.B.A.

[25] The sample sizes for correlations between ranking years 2007-2008, 2008-2010, 2010-2011, 2011-2012, and 2012-2013 equaled 42, 43, 45, 44, and 47, respectively.

Table 13

*Spearman Rank Order Correlations Between The Financial Times Rankings Across Years (2000 – 2013)*

| Variables | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2000 | 1.00 | | | | | | | | | | | | | |
| 2001 | .96 | 1.00 | | | | | | | | | | | | |
| 2002 | .91 | .93 | 1.00 | | | | | | | | | | | |
| 2003 | .91 | .93 | .93 | 1.00 | | | | | | | | | | |
| 2004 | .86 | .89 | .94 | .94 | 1.00 | | | | | | | | | |
| 2005 | .84 | .84 | .84 | .87 | .85 | 1.00 | | | | | | | | |
| 2006 | .76 | .75 | .81 | .81 | .86 | .90 | 1.00 | | | | | | | |
| 2007 | .84 | .78 | .79 | .79 | .82 | .89 | .89 | 1.00 | | | | | | |
| 2008 | .86 | .82 | .81 | .82 | .83 | .79 | .83 | .88 | 1.00 | | | | | |
| 2009 | .87 | .84 | .84 | .82 | .85 | .85 | .85 | .86 | .87 | 1.00 | | | | |
| 2010 | .89 | .89 | .88 | .90 | .91 | .85 | .84 | .81 | .83 | .89 | 1.00 | | | |
| 2011 | .91 | .88 | .85 | .87 | .88 | .89 | .82 | .87 | .82 | .90 | .92 | 1.00 | | |
| 2012 | .90 | .86 | .84 | .87 | .85 | .82 | .78 | .81 | .83 | .88 | .94 | .92 | 1.00 | |
| 2013 | .92 | .89 | .85 | .89 | .84 | .86 | .80 | .85 | .84 | .90 | .90 | .92 | .92 | 1.00 |

*Note.* All values are statistically significant, $p < .001$.

programs, ranging from 40 to 68 programs. As shown in Table 13, the lag 1 correlations[26] varied in magnitude from .85 to .96, with a weighted average correlation of .91. As a result, on average, 83% of the variance in the current *Financial Times* ranking was shared with the previous year's ranking.

Table 14 presents the Spearman rank order correlations between business school rankings published in the six issues of *Forbes* between 2000 and 2013. The correlations were computed using pairwise deletion for the sample sizes of M.B.A. programs, ranging from 44 to 83 programs. As demonstrated in Table 14, the lag 1 correlations[27] ranged in magnitude from .87 to .95, with a weighted average correlation of .94. As a result, on average, 88% of the variance in the current *Forbes* ranking was shared with the previous year's ranking.

Table 15 shows the Spearman rank order correlations between business school rankings published in the 14 issues of *U.S. News & World Report* between 2000 and 2013. The correlations were calculated using pairwise deletion for the sample sizes of M.B.A. programs, ranging from 44 to 79 programs. As indicated in Table 15, the lag 1

---

[26] The sample sizes for correlations between ranking years 2000-2001, 2001-2002, 2002-2003, 2003-2004, 2004-2005, 2005-2006, 2006-2007, 2007-2008, 2008-2009, 2009-2010, 2010-2011, 2011-2012, and 2012-2013 equaled 43, 50, 54, 54, 53, 52, 54, 56, 52, 54, 48, 47, and 46, respectively.

[27] The sample sizes for correlations between ranking years 2003-2005, 2005-2007, 2007-2009, 2009-2011, and 2011-2013 equaled 47, 52, 56, 70, and 67, respectively.

Table 14

*Spearman Rank Order Correlations Between Forbes Rankings Across Years (2003 – 2013)*

| Variables | 2003 | 2005 | 2007 | 2009 | 2011 | 2013 |
|---|---|---|---|---|---|---|
| 2003 | 1.00 | | | | | |
| 2005 | .87 | 1.00 | | | | |
| 2007 | .82 | .91 | 1.00 | | | |
| 2009 | .83 | .92 | .96 | 1.00 | | |
| 2011 | .81 | .90 | .95 | .97 | 1.00 | |
| 2013 | .80 | .88 | .91 | .92 | .95 | 1.00 |

*Note.* All values are statistically significant, $p < .001$.

correlations[28] varied in magnitude from .90 to .98, with a weighted average correlation of .95. Therefore, on average, 90% of the variance in the current *U.S. News & World Report* ranking was shared with the previous year's ranking.

The results of the correlational analyses of the five ranking systems of full-time M.B.A. programs showed considerable congruence among the five weighted average correlations, ranging from .91 to .96. Consequently, on average, for *Businessweek, The Economist, The Financial Times, Forbes,* and *U.S. News & World Report*, between 83% and 92% of the variance in the current year's ranking was shared with the previous year's

[28] The sample sizes for correlations between ranking years 2000-2001, 2001-2002, 2002-2003, 2003-2004, 2004-2005, 2005-2006, 2006-2007, 2007-2008, 2008-2009, 2009-2010, 2010-2011, 2011-2012, and 2012-2013 equaled 48, 50, 47, 47, 49, 47, 48, 48, 48, 49, 51, 67, and 67, respectively.

Table 15

*Spearman Rank Order Correlations Between U.S. News & World Report Rankings Across Years (2000 – 2013)*

| Variables | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2000 | 1.00 | | | | | | | | | | | | | |
| 2001 | .96 | 1.00 | | | | | | | | | | | | |
| 2002 | .95 | .94 | 1.00 | | | | | | | | | | | |
| 2003 | .89 | .89 | .90 | 1.00 | | | | | | | | | | |
| 2004 | .86 | .87 | .84 | .90 | 1.00 | | | | | | | | | |
| 2005 | .90 | .90 | .85 | .90 | .94 | 1.00 | | | | | | | | |
| 2006 | .90 | .91 | .88 | .92 | .94 | .96 | 1.00 | | | | | | | |
| 2007 | .93 | .93 | .87 | .89 | .92 | .94 | .94 | 1.00 | | | | | | |
| 2008 | .91 | .91 | .89 | .91 | .90 | .92 | .93 | .92 | 1.00 | | | | | |
| 2009 | .89 | .91 | .88 | .91 | .89 | .91 | .91 | .93 | .97 | 1.00 | | | | |
| 2010 | .90 | .90 | .89 | .89 | .85 | .88 | .89 | .90 | .96 | .96 | 1.00 | | | |
| 2011 | .87 | .86 | .85 | .89 | .84 | .84 | .85 | .87 | .93 | .92 | .95 | 1.00 | | |
| 2012 | .87 | .90 | .87 | .91 | .87 | .86 | .87 | .89 | .94 | .95 | .95 | .98 | 1.00 | |
| 2013 | .88 | .90 | .86 | .91 | .87 | .89 | .89 | .90 | .94 | .96 | .95 | .95 | .97 | 1.00 |

*Note.* All values are statistically significant, $p < .001$.

ranking. Thus, the five ranking systems of full-time M.B.A. programs in the U.S. demonstrated consistent statistical results with regard to their stability over time and reliability.

**Construct validity of the full-time M.B.A. rankings in the United States: Estimating construct validity with CFA of multitrait-multimethod matrices.**

***Overall model specification.*** As it was assumed due to the nature of the data – rank-ordered data[29], the assumption of multivariate normality of the data was violated. Table 16 presents the results for the four multivariate normality tests, including Mardia's, Henze-Zirkler's, Royston's Multivariate Normality Tests, and the Shapiro-Francia Test for Normality. The results demonstrated that both multivariate skewness ($p < .001$) and multivariate kurtosis ($p < .001$) values obtained from the Mardia's test violated multivariate normality. In addition, according to the Henze-Zirkler's, Royston's Multivariate Normality Tests, and the Shapiro-Francia Test for Normality, the data set did not appear to follow a multivariate normal distribution, with $p < .0001$, $p = .040$, and

[29] Although the data were ranked-ordered, PRELIS 9.10 (Jöreskog & Sörbom, 2014), the statistical software used for the analysis of univariate normality and part of the analysis of multivariate normality, treated the data as continuous due to the large number of categories in each variable, producing the following warning message for all nine variables "WARNING: *variable name* has more than 15 categories and will be treated as continuous". Therefore, I evaluated the data for multivariate and univariate normality.

$p < .001$, respectively. At the same time, relative multivariate kurtosis value equaled 1.104, which means that no serious violations of normality existed[30].

Table 16

*Summary of Results of Multivariate Normality Tests*

| Multivariate Normality Test | | Estimate | *p* |
|---|---|---|---|
| Mardia's | Multivariate skewness | 36.02 | |
| | Chi-square value of skewness statistic | 252.11 | < .001 |
| | Multivariate kurtosis | 109.34 | |
| | z value of kurtosis estimate | 2.38 | < .001 |
| Henze-Zirkler's | Henze-Zirkler statistic | 1.17 | < .001 |
| Royston's | Royston's statistic | 13.52 | .040 |
| Shapiro-Francia | Shapiro-Francia statistic | 0.7781 | < .001 |

In this respect, Kline (2011) warned against relying on the results of multivariate normality tests solely and recommended exploring univariate normality of single variables. Preliminary analyses demonstrated that absolute values of univariate skeweness and kurtosis for all variables were below the recommended values of 3 and 10, respectively, for structural equation modeling (Kline, 2011). The values of skewness ranged from .05 to .48. The values of kurtosis ranged from -.47 to -1.2. The data were

[30] For structural equation modeling, Kline (2011) recommended that absolute values of univariate skeweness and kurtosis of 3 and 10, respectively, be the cutoff points indicating serious violations of normality.

also evaluated for univariate outliers by examining distance, leverage, and influence (Stevens, 1984). The cutoff points for measures of distance were values less than -3 and greater than +3. Leverage values that were three times greater than the sample average leverage value were considered high. Finally, values of Cook's D greater than 1 were also considered high. Based on the values of distance, leverage, and influence calculated for each of the nine variables, no univariate influential points were detected.

The data were also examined for multicollinearity. For all the six multitrait-multioccasion models, the condition number, proposed by Belsley, Kuh, and Welsh (1980) to detect multicollinearity, equaled 11.10. This value was much smaller than 30, which is a cutoff point indicating serious multicolinearity according to Belsley et al. (1980). The R version 3.1.0 (2014-04-10) and PRELIS 9.10 (Jöreskog & Sörbom, 2014) statistical software were used to screen the data for multivariate and univariate normality and multicollinearity.

Because the data were in violation of the assumption of multivariate normality, the choice of the method of estimation presented a particular concern in this research. A wealth of studies has examined the performance of continuous and/or categorical estimation methods with non-normal continuous and ordinal data in structural equation modeling. There is mixed evidence on the effect of non-normality on model fit indexes, parameter estimates, and standard errors of parameters depending on the method of estimation and sample size. The most widely recommended estimation methods for continuous non-normal and/or categorical variables include robust weighted least squares (other names of RWLS include diagonally weighted least squares (DLWS) and weighted

least squares means and variance adjusted (WLSMV) estimation) (Beauducel, & Herzberg, 2006; Brown, 2006; Flora & Curran, 2004; Hox, Maas, & Brinkhuis, 2010) and robust maximum likelihood (RLM) (Brown, 2006). Weighted least squares (WLS), diagonally weighted least squares (DWLS), robust maximum likelihood (RML), and the method of normal scores are available to fit structural equation models to non-normal data in LISREL (Jöreskog, Sörbom, Du Toit, & Du Toit, 2001). At the same time, Hox, Maas, and Brinkhuis (2010) warned against using the robust methods with small sample sizes. In contrast to recommendations in the studies discussed above, some studies demonstrated that maximum likelihood estimation was not inferior to other methods in terms of parameter bias and model fit under conditions of misspecification and non-normality (Luo, 2011; Olsson, Foss, Troye, & Howell, 2000). In the same vein, Beauducel and Herzberg (2006) claimed that with categorical data robust methods like weighted least squares means and variances adjusted (WLSMV) estimation "cannot outperform ML estimation when the number of categories is very large" (p. 201). Taking into account that the number of categories in the rank-ordered data is large and the sample size is very small (42 business schools) in this study, I accepted the recommendation of Iacobucci (2010) to use maximum likelihood estimation. Table 17 contains the correlation matrix, standard deviations, and means for *Businessweek* indicators. The R version 3.1.0 (2014-04-10) was used to calculate Spearman rank order correlations between *Businessweek* rankings indicators, their standard deviations, and means.

Table 17

*Summary of Spearman Rank Order Correlations Between Businessweek Indicators, Standard Deviations, and Means*

| Variables | Grad Poll M1 | Grad Poll M2 | Grad Poll M3 | Corp Poll M1 | Corp Poll M2 | Corp Poll M3 | Intel Capital M1 | Intel Capital M2 | Intel Capital M3 |
|---|---|---|---|---|---|---|---|---|---|
| Grad Poll M1 | 1.00 | | | | | | | | |
| Grad Poll M2 | .86 | 1.00 | | | | | | | |
| Grad Poll M3 | .79 | .75 | 1.00 | | | | | | |
| Corp Poll M1 | .81 | .74 | .57 | 1.00 | | | | | |
| Corp Poll M2 | .73 | .63 | .48 | .89 | 1.00 | | | | |
| Corp Poll M3 | .81 | .72 | .61 | .88 | .86 | 1.00 | | | |
| Intel Capital M1 | .63 | .51 | .59 | .48 | .47 | .61 | 1.00 | | |
| Intel Capital M2 | .62 | .58 | .62 | .51 | .47 | .64 | .92 | 1.00 | |
| Intel Capital M3 | .59 | .56 | .54 | .45 | .44 | .57 | .85 | .83 | 1.00 |
| *SD* | 12.47 | 14.09 | 14.87 | 12.57 | 15.37 | 14.69 | 13.30 | 15.73 | 16.51 |
| *M* | 21.62 | 22.71 | 23.38 | 21.71 | 25.24 | 23.05 | 22.48 | 26.12 | 25.33 |

***Results.*** Table 18 presents goodness-of fit indexes for the six multitrait-multioccasion, confirmatory factor analysis models proposed to assess the degree of convergent and discriminant validity of the *Bloomberg Businessweek* rankings. Following Jaccard and Wan's (1996) recommendation to use goodness-of-fit indexes from different classes in order to avoid limitations of each index, this study utilized both absolute and incremental fit indexes to measure model fit[31]. The values of goodness-of-fit indexes indicated in Table 18 demonstrate that the correlated traits, correlated method

---

[31] The chi-square test of model fit "assesses the magnitude of discrepancy between the sample and fitted covariance matrices" (Hu & Bentler, 1999, p. 2). An insignificant *p*-value ($p > .05$) for low chi-square statistics relative to degrees of freedom indicates a good model fit. Root Mean Square Error of Approximation (RMSEA) is expected to be less than .05 for a good model fit (Steiger, 1990), although the upper limit of .07 (Steiger, 2007) is also accepted. Values of Standardized Root Mean Square Residual (SRMR) that are less than .05 indicate a good model fit (Byrne, 1998), nevertheless a cutoff value of .08 (Hu & Bentler, 1999) is also acceptable. Finally, Goodness of Fit Index (GFI) should exceed .90 (Hu & Bentler, 1999). With regard to the indexes that compare the target and the baseline models, a model is considered acceptable if the Comparative Fit Index (CFI) exceeds .93 (Byrne, 1994) or a more conservative estimate of .95 (Hu & Bentler, 1999); Normed Fit Index (NFI) exceeds .90 (Byrne, 1994) or .95 (Hu & Bentler, 1999; Schumacker & Lomax, 2004); and Non-Normed Fit Index (or Tucker Lewis index) has value over .90 or even .95 (Hu & Bentler, 1999). Goodness of Fit Index (GFI) should exceed .90 (Hu & Bentler, 1999).

Table 18

*Goodness-of-Fit Indexes for Nested Sequence of Multitrait-Multioccasion, Confirmatory Factor Analysis Models*

| Model | $\chi^2$ | *p* | *df* | CFI | NFI | NNFI | RMSEA | PCLOSE | SRMR | GFI |
|---|---|---|---|---|---|---|---|---|---|---|
| 1. Correlated Traits, Correlated Methods | 10.46 | .575 | 12 | 1.00 | .98 | 1.01 | .00 | .66 | .05 | .95 |
| 2. No Traits, Correlated Methods | 142.78 | < .001 | 24 | .81 | .79 | .72 | .34 | < .001 | .11 | .57 |
| 3. Orthogonal Traits, Correlated Methods | 66.64 | < .001 | 15 | .92 | .90 | .80 | .29 | < .001 | .44 | .76 |
| 4. Correlated Traits, Orthogonal Methods | 22.72 | .090 | 15 | .99 | .97 | .97 | .11 | .15 | .053 | .90 |
| 5. Correlated Traits, Correlated Uniqueness | 22.12 | .105 | 15 | .99 | .97 | .97 | .11 | .17 | .053 | .90 |
| 6. Orthogonal Traits, Correlated Uniqueness | 87.66 | < .001 | 18 | .89 | .87 | .78 | .30 | < .001 | .47 | .74 |

*Note.* CFI = Comparative Fit Index; NFI = Normed Fit Index (or Bentler-Bonett index); NNFI = Non-Normed Fit Index (or Tucker Lewis index); RMSEA = Root Mean Square Error of Approximation; PCLOSE = p-value for test of close fit (RMSEA < .05); SRMR = Standardized Root Mean Square Residual; GFI = Goodness of Fit Index.

model was the best fitting model among the four correlated methods CFA models, while the correlated traits, correlated uniqueness model proved to have better fit than the orthogonal traits, correlated uniqueness model.

The statistically significant difference between the fit of Models 1 and 2 was supportive of the convergent validity of the conceptualization of the *Bloomberg Businessweek* ranking system presented in Table 19. The existence of traits in Model 1 significantly improved the fit, and the measures of the same trait obtained by different methods converged. The statistically significant difference between the fit of Models 1 and 3 provided evidence of the discriminant validity of the traits. When the three traits representing different facets of the business school quality were allowed to correlate freely, the fit of the model significantly improved, which shows that the three business school quality traits were not orthogonal and diverged meaningfully. Finally, discriminant validity of time-specific methods, represented by the *Bloomberg Businessweek* ranking instruments in 2008, 2010, and 2012, was supported by the significant difference between the fit of Models 1 and 4. When the model distinguished between time-specific methods, model fit improved significantly, which demonstrates that each of the three time-specific methods employed provided nonredundant information about the traits that form the overarching construct of the business school quality.

In this study, assessment of construct validity was based on the premise that "when the results indicate that convergent and discriminant validity are high and method effects are negligible, construct validity is supported" (Brown, 2006, p. 217). The results

Table 19

*Differential Goodness-of-Fit Indexes for Multitrait-Multioccasion, Confirmatory Factor Analysis Nested Model Comparisons*

| Model comparisons | Δχ | Δ*df* | *p* | ΔCFI | ΔNFI | ΔNNFI | ΔRMSEA | Δ SRMR | ΔGFI |
|---|---|---|---|---|---|---|---|---|---|
| Test of convergent validity | | | | | | | | | |
| Model 1 vs. Model 2 (traits) | 132.32 | 12 | < .001 | .19 | .19 | .29 | -.34 | -.06 | .38 |
| Test of discriminant validity | | | | | | | | | |
| Model 1 vs. Model 3 (traits) | 56.18 | 3 | < .001 | .08 | .08 | .21 | -.29 | -.39 | .19 |
| Model 1 vs. Model 4 (methods) | 12.26 | 3 | 0.007 | .01 | .01 | .04 | -.11 | -.003 | .05 |
| Model 5 vs. Model 6 (traits) | 65.54 | 3 | < .001 | .10 | .10 | .19 | -.19 | -.417 | .16 |

*Note.* CFI = Comparative Fit Index; NFI = Normed Fit Index (or Bentler-Bonett index); NNFI = Non-Normed Fit Index (or Tucker Lewis index); RMSEA = Root Mean Square Error of Approximation; SRMR = Standardized Root Mean Square Residual; GFI = Goodness of Fit Index.

in Tables 20 and 21 suggest the construct validity of the *Bloomberg Businessweek* ranking system by showing adequate convergent validity and discriminant validity and the absence of method effects. The results came from the two best fitting models, one model came from a set of correlated methods models, and the other model came from a set of correlated uniqueness models, which triangulated the findings of the study.

As demonstrated in Tables 20 and 21, the trait factor loadings were statistically significant and consistently large, ranging from .75 to .99 in the correlated traits, correlated methods model and from .80 to .98 in the correlated traits, correlated uniqueness model. These findings suggest that the measures were reasonable indicators of their respective trait factors and could be viewed in support of convergent validity.

In the correlated traits, correlated methods model, moderate correlations among trait factors (.55 and .63), except the high correlation between Graduate Poll and Corporate Poll factors (.83), were indicative of good discriminant validity with one exception: Graduate Poll and Corporate Poll factors did not represent entirely different traits, as evidenced by the high estimate of their correlation. In the same fashion, correlations among trait factors in the correlated traits, correlated uniqueness model were not high, ranging from .58 to .65, with the exception of the high correlation between Graduate Poll and Corporate Poll factors (.85). The latter may again indicate that the traits Graduate Poll and Corporate Poll were not distinct. As *Bloomberg Businessweek* is

Table 20

*Completely Standardized Estimates for the Correlated Traits, Correlated Methods Model of the Multitrait-Multioccasion Matrix*

| | Trait factor loadings | | | Method factor loadings | | |
|---|---|---|---|---|---|---|
| | Graduate Poll | Corporate Poll | Intellectual Capital | Method 1 | Method 2 | Method 3 |
| Graduate Poll Method 1 | .99* | | | .10 | | |
| Graduate Poll Method 2 | .86* | | | | .16 | |
| Graduate Poll Method 3 | .77* | | | | | .26 |
| Corporate Poll Method 1 | | .97* | | -.07 | | |
| Corporate Poll Method 2 | | .92* | | | -.02 | |
| Corporate Poll Method 3 | | .94* | | | | .26 |
| Intellectual Capital Method 1 | | | .96* | .20 | | |
| Intellectual Capital Method 2 | | | .92* | | .23 | |
| Intellectual Capital Method 3 | | | .75* | | | .25 |
| | Trait factor correlations | | | | | |
| | Graduate Poll | Corporate Poll | Intellectual Capital | | | |
| Graduate Poll | 1.00 | | | | | |
| Corporate Poll | .83* | 1.00 | | | | |
| Intellectual Capital | .63* | .55* | 1.00 | | | |

*Note.* $*p < .05$

Table 21

*Completely Standardized Estimates for the Correlated Traits, Correlated Uniqueness Model of the Multitrait-Multioccasion Matrix*

| | Trait factor loadings | | | Correlated uniqueness | | |
|---|---|---|---|---|---|---|
| | Graduate Poll | Corporate Poll | Intellectual Capital | | | |
| Graduate Poll Method 1 | .98* | | | 1.00 | | |
| Graduate Poll Method 2 | .89* | | | 1.00 | | |
| Graduate Poll Method 3 | .80* | | | 1.00 | | |
| Corporate Poll Method 1 | | .95* | | .01 | 1.00 | |
| Corporate Poll Method 2 | | .93* | | -.04 | 1.00 | |
| Corporate Poll Method 3 | | .93* | | .03 | 1.00 | |
| Intellectual Capital Method 1 | | | .97* | .00 | .04 | 1.00 |
| Intellectual Capital Method 2 | | | .95* | -.03 | .03 | 1.00 |
| Intellectual Capital Method 3 | | | .88* | .00 | -.03 | 1.00 |
| | Trait factor correlations | | | | | |
| | Graduate Poll | Corporate Poll | Intellectual Capital | | | |
| Graduate Poll | 1.00 | | | | | |
| Corporate Poll | .85* | 1.00 | | | | |
| Intellectual Capital | .65* | .58* | 1.00 | | | |

*Note.* $*p < .05$.

protective of the instruments[32] it uses to gauge the opinions of business school graduates and corporate recruiters, the finding about the unclear distinction between Graduate Poll and Corporate Poll factors could be attributed to the measurement error or sampling problems in the process of data collection.

Finally, small or nonsignificant method factor loadings implied an absence of methods effects. In the correlated traits, correlated methods model, the method factor loadings were consistently small, ranging from -.07 to .26, and were not statistically significant, which indicated the absence of the time-specific method effects. In the correlated traits, correlated uniqueness model, the correlated uniqunesses among indicators assessed by the same method were also consistently small, ranging from -.04 to .04, and were not statistically significant, which pointed to the absence of time-specific method biases.

**Research Question 3**

**Conundrums of reputation: Anchoring effects in the full-time M.B.A. rankings in the United States.**

***Overall model specification.*** Prior to the path analysis, the endogenous continuous variables were evaluated for univariate and multivariate normality. The R version 3.1.0 (2014-04-10) and PRELIS 9.10 (Jöreskog & Sörbom, 2014) statistical software were used to screen the data for multivariate and univariate normality. Table 22

---

[32] In 2004, *Bloomberg Businessweek* published 15 key questions out of 45 total in its survey of M.B.A. graduates. The author of the study could not find any publications of the survey, which *Bloomberg Businessweek* administers to recruiters.

presents the results for the four multivariate normality tests, including Mardia's, Henze-Zirkler's, Royston's Multivariate Normality Tests, and the Shapiro-Francia Test for Normality for the three models under analysis. The results for Model 1 demonstrated that both multivariate skewness ($p = .133$) and multivariate kurtosis ($p = .895$) values obtained from the Mardia's test did not violate multivariate normality. However, according to the Henze-Zirkler's, Royston's Multivariate Normality Tests, and the Shapiro-Francia Test for Normality, the data set did not appear to follow a multivariate normal distribution, with $p = .009$, $p = .0002$, and $p = .006$, respectively. At the same time, the relative multivariate kurtosis value equaled 1.012, which means that no serious violations of normality existed.

For Model 2, the results from the Mardia's test demonstrated that multivariate skewness ($p = .131$) did not violate multivariate normality, while multivariate kurtosis ($p = .032$) was in violation with this assumption. The results of the Shapiro-Francia Test for Normality supported multivariate normality of the data ($p = .371$). By contrast, the results from the Henze-Zirkler's and Royston's Multivariate Normality Tests showed that the data were not multivariate normal, with $p < .001$ and $p = .0002$, respectively. For Model 2, the relative multivariate kurtosis value equaled 0.804, which indicated no serious violations of normality.

For Model 3, the results from the Mardia's test demonstrated that multivariate skewness ($p = .169$) did not violate multivariate normality, while multivariate kurtosis ($p = .040$) did. The test statistics obtained from the Shapiro-Francia Test for Normality did not show departures from multivariate normality ($p = .408$), while the results from the

Table 22

*Summary of Results of Multivariate Normality Tests for Model 1, Model 2, and Model 3*

| Multivariate Normality Test | | Model 1 (2011-2012) | | Model 2 (2012-2013) | | Model 3 (2011-2013) | |
|---|---|---|---|---|---|---|---|
| | | Estimate | *p* | Estimate | *p* | Estimate | *p* |
| Mardia's | Multivariate skewness | 1.40 | | 1.41 | | 1.32 | |
| | Chi-square value of skewness statistic | 14.97 | .133 | 15.04 | .131 | 14.09 | .169 |
| | Multivariate kurtosis | 15.18 | | 12.08 | | 12.18 | |
| | z value of kurtosis estimate | 0.13 | .895 | -2.14 | .032 | -2.06 | .040 |
| Henze-Zirkler's | Henze-Zirkler statistic | 1.10 | .009 | 1.48 | < .001 | 1.41 | .0002 |
| Royston's | Royston's statistic | 17.01 | .0002 | 16.76 | .0002 | 17.12 | .0002 |
| Shapiro-Francia | Shapiro-Francia statistic | 0.9424 | .006 | 0.9813 | .371 | 0.9822 | .408 |

Henze-Zirkler's and Royston's Multivariate Normality Tests indicated the opposite, with $p = .0002$ and $p = .0002$, respectively. The relative multivariate kurtosis value equaled 0.812, which showed no serious violations of normality. Thus, I obtained mixed results with regard to the assumption of multivariate normality of the endogenous continuous variables used in Models 1, 2, and 3.

As Kline (2011) cautioned against relying on the results of multivariate normality tests solely and recommended exploring univariate normality of single variables, I examined skewness and kurtosis statistics for each variable in the three models. Preliminary analyses demonstrated that absolute values of univariate skeweness and kurtosis for all variables were below 1, which was much smaller than the recommended values of 3 and 10, respectively, for structural equation modeling (Kline, 2011). The data were also evaluated for univariate outliers by examining distance, leverage, and influence (Stevens, 1984). The cutoff points for measures of distance were values less than -3 and greater than +3. Leverage values that were three times greater than the sample average leverage value were considered high. Finally, values of Cook's D greater than 1 were also considered high. Based on the values of distance, leverage, and influence calculated for each of the nine variables, no univariate influential outliers were detected.

As linearity and homoscedacity among residuals are aspects of multivariate normality (Kline, 2011), the endogenous data were screened for violations of assumptions of linearity and homogeneity of variance. I examined all bivariate scatterplots of the independent (exogenous) and the dependent variables from multiple regression models fitted to endogenous variables from Model 1, 2, and 3. Examination of

the bivariate scatterplots suggested that a linear model was plausible for the relationship between the endogenous and exogenous variables. The assumption of homoscedacity was tested with Levene's test for homogeneity of variance (Levene, 1960). The assumption of homoscedacity was met in all cases.

The data were also examined for multicollinearity in the three path analysis models. The condition number, proposed by Belsley et al. (1980) to detect multicollinearity, equaled 10.83 for Model 1, 11.16 for Model 2, and 11.15 for Model 3. All the three values were much smaller than 30, which is a cutoff point indicating serious multicolinearity according to Belsley et al. (1980).

Tables 23, 24, and 25 present the correlation matrices and standard deviations for the three models. The R version 3.1.0 (2014-04-10) was used to calculate Pearson correlations and standard deviations.

***Results.*** All the three structural equation models were saturated, $\chi^2$ (0) = 0.00, $p$ = 1.00, which means that the model fit was perfect. Tables 26, 27, and 28 present the results of path analysis for Model 1, Model 2, and Model 3. Figures 9, 10, and 11 show only the direct effects for variables in Model 1, Model 2, and Model 3, respectively. As hypothesized, the current university reputation had a strong direct effect on the current university rank and remained the same in both 2011 and 2012, $\beta$ = .93, $p$ < .05. As demonstrated in Figures 9 and 10, the current business school reputation was strongly related to the current business school rank, gradually increasing with each subsequent year, $\beta$ = .88, $p$ < .05 in 2012 and $\beta$ = .91, $p$ < .05 in 2013. As predicted, the current university reputation contributed positively to the future reputation of its business school,

Table 23

*Summary of Pearson Correlations and Standard Deviations for Model 1*

| Variables | University reputation 2011 | University rank 2011 | B-school reputation 2012 | B-school rank 2012 |
|---|---|---|---|---|
| University reputation 2011 | 1.00 | | | |
| University rank 2011 | .93 | 1.00 | | |
| B-school reputation 2012 | .88 | .82 | 1.00 | |
| B-school rank 2012 | .87 | .82 | .97 | 1.00 |
| *SD* | 9.96 | 16.30 | 0.57 | 15.38 |

Table 24

*Summary of Pearson Correlations and Standard Deviations for Model 2*

| Variables | University reputation 2012 | University rank 2012 | B-school reputation 2013 | B-school rank 2013 |
|---|---|---|---|---|
| University reputation 2012 | 1.00 | | | |
| University rank 2012 | .93 | 1.00 | | |
| B-school reputation 2013 | .87 | .79 | 1.00 | |
| B-school rank 2013 | .86 | .80 | .97 | 1.00 |
| *SD* | 9.87 | 16.90 | 0.60 | 15.91 |

Table 25

*Summary of Pearson Correlations and Standard Deviations for Model 3*

| Variables | University reputation 2011 | University rank 2011 | B-school reputation 2013 | B-school rank 2013 |
|---|---|---|---|---|
| University reputation 2011 | 1.00 | | | |
| University rank 2011 | .93 | 1.00 | | |
| B-school reputation 2013 | .87 | .80 | 1.00 | |
| B-school rank 2013 | .86 | .80 | .97 | 1.00 |
| *SD* | 9.96 | 16.30 | 0.60 | 15.91 |

Table 26

*Completely Standardized Estimates for Direct, Indirect, and Total Effects for Structural Equation Model 1*

| | University reputation 2011 | University rank 2011 | B-school reputation 2012 | $R^2$ |
|---|---|---|---|---|
| Direct | | | | |
| University rank 2011 | .93* | | | .86 |
| B-school reputation 2012 | .89* | -.01 | | .77 |
| B-school rank 2012 | .03 | .07 | .88* | .94 |
| Indirect | | | | |
| University rank 2011 | | | | |
| B-school reputation 2012 | -.01 | | | |
| B-school rank 2012 | .84* | -.01 | | |

(continued)

Table 26 (continued)

| | University reputation 2011 | University rank 2011 | B-school reputation 2012 | $R^2$ |
|---|---|---|---|---|
| Total | | | | |
| University rank 2011 | .93* | | | |
| B-school reputation 2012 | .88* | -.01 | | |
| B-school rank 2012 | .87* | .06 | .88* | |

*Note.* *$p < .05$.

Table 27

*Completely Standardized Estimates for Direct, Indirect, and Total Effects for Structural Equation Model 2*

| | University reputation 2012 | University rank 2012 | B-school reputation 2013 | $R^2$ |
|---|---|---|---|---|
| Direct | | | | |
| University rank 2012 | .93* | | | .86 |
| B-school reputation 2013 | .96* | -.11 | | .75 |
| B-school rank 2013 | -.01 | .09 | .91* | .94 |
| Indirect | | | | |
| University rank 2012 | | | | |
| B-school reputation 2013 | -.10 | | | |
| B-school rank 2013 | .87* | -.10 | | |
| Total | | | | |
| University rank 2012 | .93* | | | |
| B-school reputation 2013 | .87* | -.11 | | |
| B-school rank 2013 | .86* | -.01 | .91* | |

*Note.* *$p < .05$.

Table 28

*Completely Standardized Estimates for Direct, Indirect, and Total Effects for Structural Equation Model 3*

| | University reputation 2011 | University rank 2011 | B-school reputation 2013 | $R^2$ |
|---|---|---|---|---|
| Direct | | | | |
| University rank 2011 | .93* | | | .86 |
| B-school reputation 2013 | .94* | -.08 | | .76 |
| B-school rank 2013 | -.00 | .07 | .92* | .94 |
| Indirect | | | | |
| University rank 2011 | | | | |
| B-school reputation 2013 | -.07 | | | |
| B-school rank 2013 | .87* | -.07 | | |
| Total | | | | |
| University rank 2011 | .93* | | | |
| B-school reputation 2013 | .87* | -.08 | | |
| B-school rank 2013 | .86* | .00 | .92* | |

*Note.* *$p < .05$.

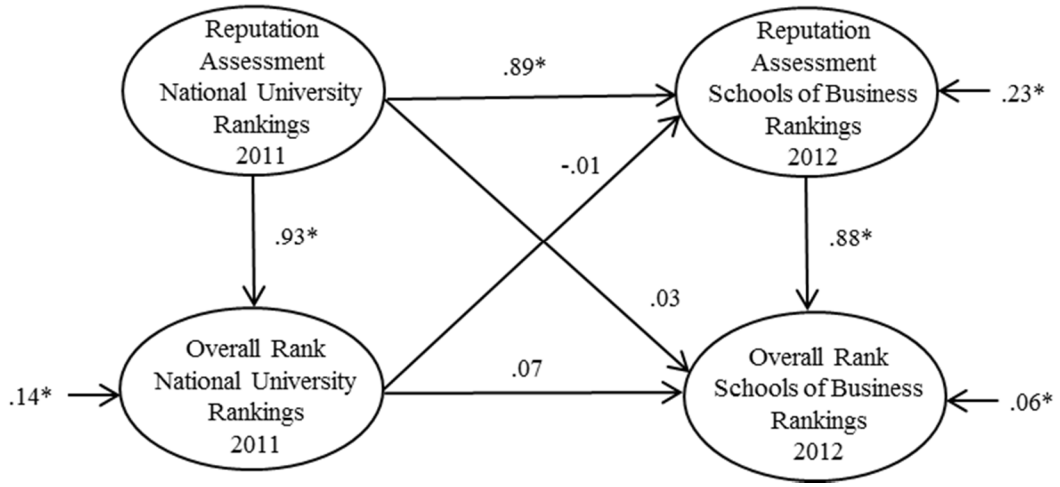


*Figure 9.* Completely standardized estimates for direct effects and endogenous disturbances for Structural Equation Model 1.

*$p < .05$.

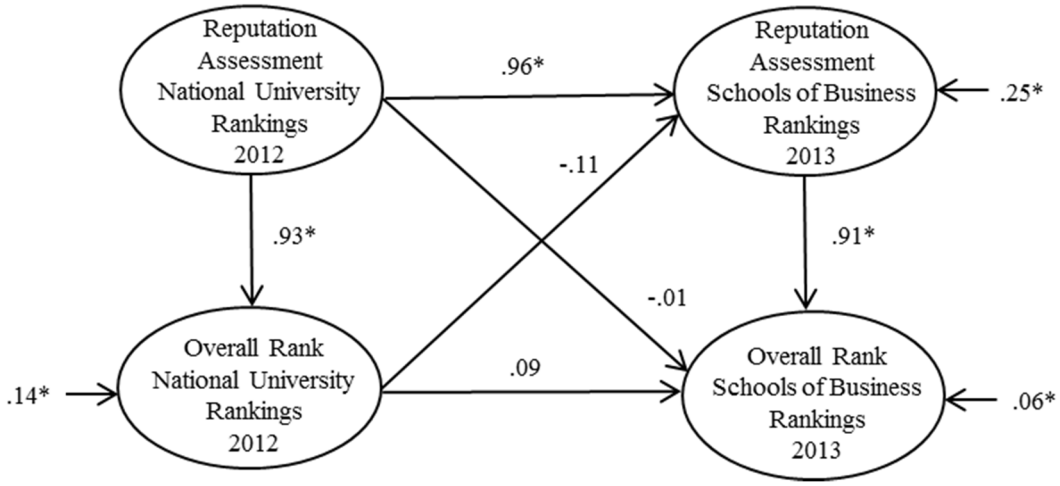


*Figure 10.* Completely standardized estimates for direct effects and endogenous disturbances for Structural Equation Model 2.

*$p < .05$.

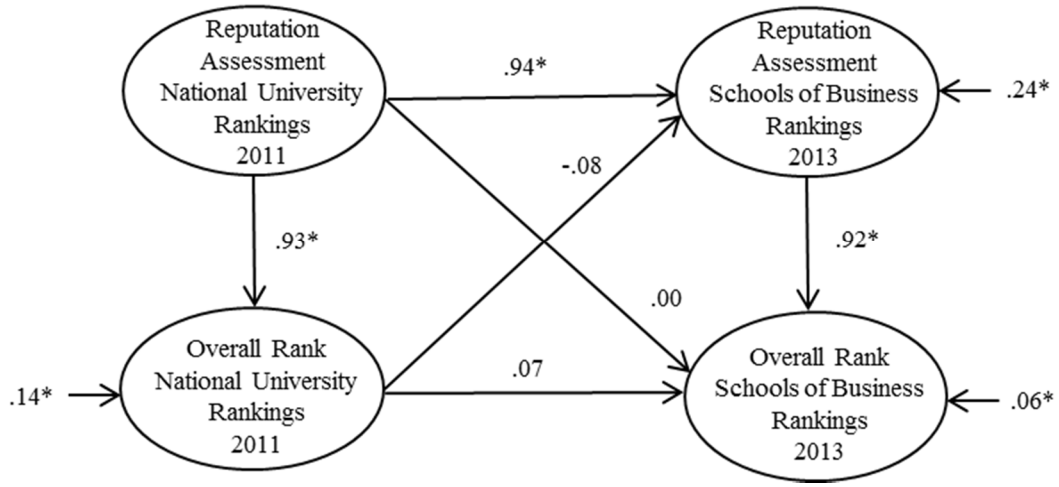


*Figure 11.* Completely standardized estimates for direct effects and endogenous disturbances for Structural Equation Model 3.

*$p < .05$.

also increasing with each subsequent year, $\beta = .89$, $p < .05$ in Model 1 and $\beta = .96$, $p < .05$ in Model 2. The current university reputation, however, was not related to the future business school rank either in 2012 or 2013. Contrary to the hypothesis, the current university rank did not have any significant effect on either future business school reputation or business school rank (see Figures 9 and 10).

The same pattern of relationship held in Model 3 that examined the effect of the university reputation and rank on the business school reputation and rank at lag 2. As shown in Figure 11, business school reputation in 2013 contributed positively to business school rank in 2013, $\beta = .92$, $p < .05$. This relationship was somewhat stronger in Model 3 than in Model 2, $\beta = .92$ compared to $\beta = .91$. In Model 3, the relationship between university reputation and university rank in 2011 remained the same as in Model 1, $\beta = .93$, $p < .05$. With respect to the lagged effects, university reputation in 2011 was

strongly related to business school reputation in 2013, with $\beta = .94$, $p < .05$. Figure 11 demonstrates that at a longer lag the positive association between the university reputation and the business school reputation increased. Lag 2 effect of the university reputation on the business school reputation in Model 3 appeared to be larger than lag 1 effect in Model 1, $\beta = .94$, $p < .05$ in Model 3 versus $\beta = .89$, $p < .05$ in Model 1. Similar to the two previous models with a shorter lag, university reputation in 2011 was not related to business school rank in 2013. As in Models 1 and 2, the hypothesis about the impact of the current university rank on both the future business school reputation and business school rank was not supported at lag 2 (see Figure 11).

**Research Question 4**

**Are we making money for the entire institution? The impact of rankings of full-time M.B.A. programs on funding streams to institutions of higher education.**

***Overall model specification.*** Prior to the path analysis, the endogenous continuous variables were evaluated for univariate normality. The R version 3.1.0 (2014-04-10) and PRELIS 9.10 (Jöreskog & Sörbom, 2014) statistical software were used to screen the data for univariate normality. As *U.S. News & World Report* ranked 51 business schools in 2010, the initial sample size equaled 51.

I examined skewness and kurtosis statistics for each endogenous variable in the six models. Preliminary analyses demonstrated that values of univariate skeweness ranged from .38 to 1.59, and values of univariate kurtosis were between 0.04 and 2.22 for endogenous variables in the four models, which was much smaller than the recommended absolute values of 3 and 10, respectively, for structural equation modeling (Kline, 2011).

Nevertheless, two variables, R&D funding from business and R&D funding from all other sources, had the values of skeweness (3.09 and 2.72) and kurtosis (12.83 and 9.18) that suggested extreme nonnormality.

The data were also evaluated for univariate outliers by examining distance, leverage, and influence (Stevens, 1984). The cutoff points for measures of distance were values less than -3 and greater than +3. Leverage values that were three times greater than the sample average leverage value were considered high. Finally, values of Cook's D greater than 1 were also considered high. Univariate influential outliers were defined based on the values of distance, leverage, and influence calculated for each of the six variables. One potentially influential observation was detected when fitting regression models to all six endogenous variables.

As linearity and homoscedacity among residuals are aspects of normality (Kline, 2011), the endogenous data were screened for violations of assumptions of linearity and homogeneity of variance. In line with Kline's (2011) recommendation to inspect the joint distribution of any pair of variables for bivariate normality, I examined all bivariate scatterplots of the independent (exogenous) and the dependent variables from regression models fitted to endogenous variables from the six models. Examination of the bivariate scatterplots suggested that a linear model was plausible for the relationship between the endogenous and exogenous variables in Models 1, 2, 3, and 5. The assumption of homoscedacity was tested with Levene's test for homogeneity of variance (Levene, 1960). The assumption of homoscedacity was met in Models 1, 2, 3, 5, and 6. Models 4 and 6 that had the endogenous variables R&D funding from business and R&D funding

from all other sources, which had extreme values of skewness and kurtosis, did not meet the assumptions of linearity (Model 4 and 6) and homoscedacity (Model 4). When the potentially influential observation was removed from the dataset, which decreased the sample size to 50 schools, linearity improved in all the six models to a greater or lesser extent, but the assumption of homogeneity of variance was still not met in Model 4. At the same time, removing the influential outlier did not change the values of skeweness and kurtosis for variables R&D funding from business and R&D funding from all other sources in Models 4 and 6 enough to meet the cutoff values of 3 and 10 specified by Kline (2011). Therefore, I used square root transformation[33] to transform R&D funding from business and R&D funding from all other sources variables. This type of data transformation was suggested by Howell (2007), Osborn (2010), and Tabachnick and Fidell (2007) if the data distribution was positively skewed. This normalizing procedure

---

[33] Some researchers, Zumel and Mount (2014) for example, recommend log base 10 transformation for monetary units because the transformed data may be easier to read in some instances as orders of 10 seem to be natural for money. Nevertheless, R&D funding from all other sources variable contained zero values. Therefore, I tried log base 10 transformation ($\log_{10}(x)$) for R&D funding from business and log base 10 transformation plus constant ($\log_{10}(x + 10)$) for R&D funding from all other sources as this variable contained zero values. These types of transformations were suggested by Howell (2007) and Tabachnick and Fidell (2007) if the data distribution was substantially positively skewed and substantially positively skewed with zero values. Nevertheless, these procedures resulted in substantially negatively skewed data distributions.

resulted in improving normality of the distributions of the two variables, improving linearity, and in equalizing the variable variances to meet the assumptions. Transformation of the data was preferred to nonparametric methods, such as robust maximum likelihood, weighted least squares, or elliptical reweighted least squares, because these methods of estimation require a large sample size (Boomsma & Hoogland, 2001).

The data were also examined for multicollinearity in the six path analysis models. The condition number, proposed by Belsley et al. (1980) to detect multicollinearity, equaled 5.50 for Model 1, 2.26 for Model 2, 2.73 for Model 3, 2.73 for Model 4, 3.59 for Model 5, and 2.05 for Model 6. All the six values were much smaller than 30, which is a cutoff point indicating serious multicolinearity according to Belsley et al. (1980).

The correlation matrices and standard deviations for the six models are provided in Tables 29, 30, 31, 32, 33, and 34. The R version 3.1.0 (2014-04-10) was used to calculate Pearson correlations and standard deviations.

***Results.*** All the six structural equation models were saturated, $\chi^2$ (0) = 0.00, $p$ = 1.00, which means that the model fit was perfect. Figures 12, 13, 14, 15, 16 and 17 show the direct effects for variables in Model 1, Model 2, Model 3, Model 4, Model 5, and Model 6, respectively. The findings provided support for four out of six research hypotheses. As hypothesized, the rank of a business school in 2010 had a significant direct effect on future resource flows from the federal, state and local government, institutional funds, and nonprofit organizations to a research university for research and development in 2011. More specifically, as shown in Figures 12 and 16, after controlling

Table 29

*Summary of Pearson Correlations and Standard Deviations for Model 1*

| Variables | *U.S. News & World Report* b-school rank 2010 | Overall R&D of a HEI 2010 | R&D of a HEI 2011 from federal government |
|---|---|---|---|
| *U.S. News & World Report* b-school rank 2010 | 1.00 | | |
| Overall R&D of a HEI 2010 | .34 | 1.00 | |
| R&D of a HEI 2011 federal government | .44 | .92 | 1.00 |
| *SD* | 13.65 | 280,681,700.00 | 205,030,300.00 |

Table 30

*Summary of Pearson Correlations and Standard Deviations for Model 2*

| Variables | *U.S. News & World Report* b-school rank 2010 | Overall R&D of a HEI 2010 | R&D of a HEI 2011 from state & local government |
|---|---|---|---|
| *U.S. News & World Report* b-school rank 2010 | 1.00 | | |
| Overall R&D of a HEI 2010 | .34 | 1.00 | |
| R&D of a HEI 2011 state & local government | -.18 | .50 | 1.00 |
| *SD* | 13.65 | 280,681,700.00 | 31,752,260.00 |

Table 31

*Summary of Pearson Correlations and Standard Deviations for Model 3*

| Variables | *U.S. News & World Report* b-school rank 2010 | Overall R&D of a HEI 2010 | R&D of a HEI 2011 from institution funds |
|---|---|---|---|
| *U.S. News & World Report* b-school rank 2010 | 1.00 | | |
| Overall R&D of a HEI 2010 | .34 | 1.00 | |
| R&D of a HEI 2011 Institution funds | -.15 | .62 | 1.00 |
| *SD* | 13.65 | 28,0681,700.00 | 65,232,250.00 |

Table 32

*Summary of Pearson Correlations and Standard Deviations for Model 4*

| Variables | *U.S. News & World Report* b-school rank 2010 | Overall R&D of a HEI 2010 | R&D of a HEI 2011 from business |
|---|---|---|---|
| *U.S. News & World Report* b-school rank 2010 | 1.00 | | |
| Overall R&D of a HEI 2010 | .34 | 1.00 | |
| R&D of a HEI 2011 business | .39 | .73 | 1.00 |
| *SD* | 13.65 | 280,681,700.00 | 2,666.73 |

Table 33

*Summary of Pearson Correlations and Standard Deviations for Model 5*

| Variables | *U.S. News & World Report* b-school rank 2010 | Overall R&D of a HEI 2010 | R&D of a HEI 2011 from nonprofit organizations |
|---|---|---|---|
| *U.S. News & World Report* b-school rank 2010 | 1.00 | | |
| Overall R&D of a HEI 2010 | .34 | 1.00 | |
| R&D of a HEI 2011 nonprofit organizations | .48 | .82 | 1.00 |
| *SD* | 13.65 | 280,681,700.00 | 28,947,430.00 |

Table 34

*Summary of Pearson Correlations and Standard Deviations for Model 6*

| Variables | *U.S. News & World Report* b-school rank 2010 | Overall R&D of a HEI 2010 | R&D of a HEI 2011 from all other sources |
|---|---|---|---|
| *U.S. News & World Report* b-school rank 2010 | 1.00 | | |
| Overall R&D of a HEI 2010 | .34 | 1.00 | |
| R&D of a HEI 2011 all other sources | .20 | .56 | 1.00 |
| *SD* | 13.65 | 280,681,700.00 | 2,110.25 |

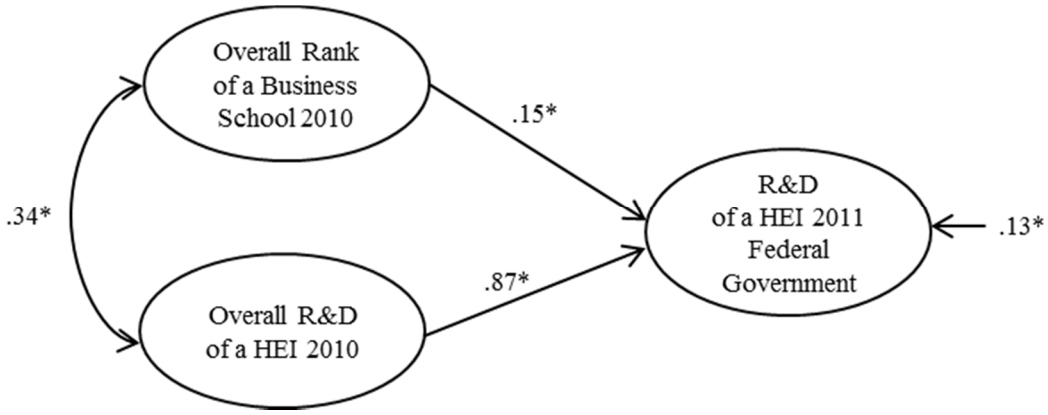


*Figure 12.* Completely standardized estimates for direct effects and endogenous disturbances for Structural Equation Model 1.

*$p < .05$.

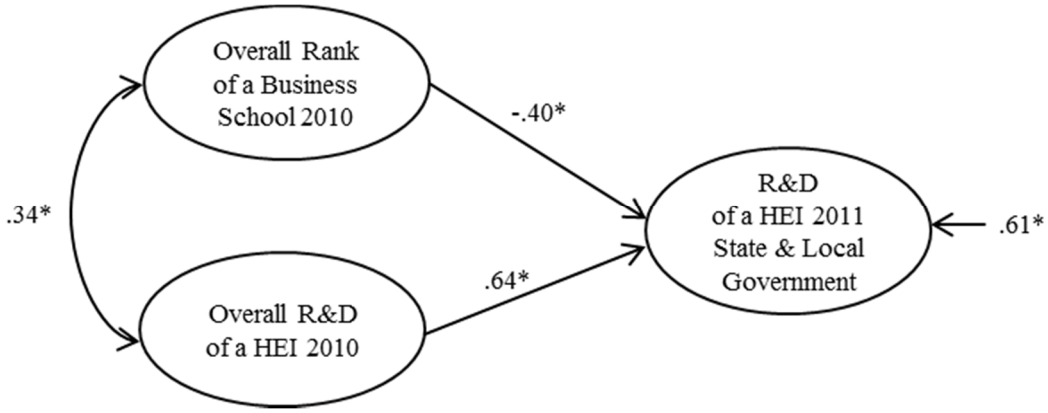


*Figure 13.* Completely standardized estimates for direct effects and endogenous disturbances for Structural Equation Model 2.

*$p < .05$.

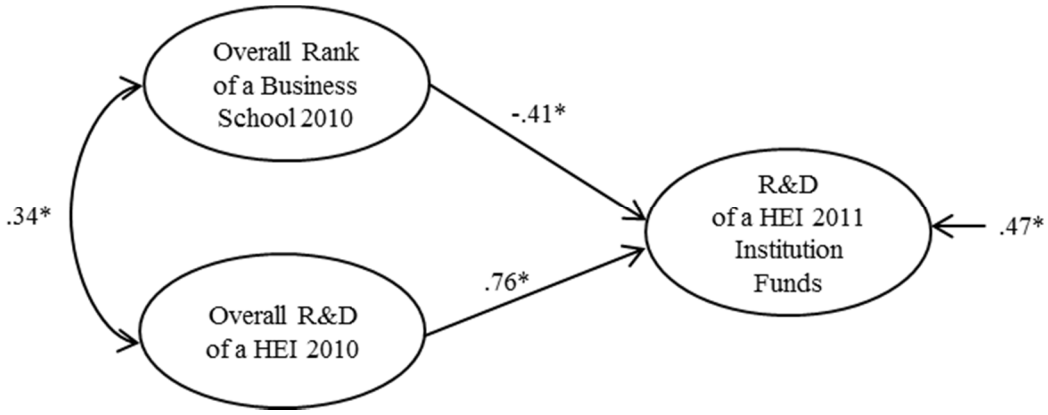


*Figure 14.* Completely standardized estimates for direct effects and endogenous disturbances for Structural Equation Model 3.

*$p < .05$.

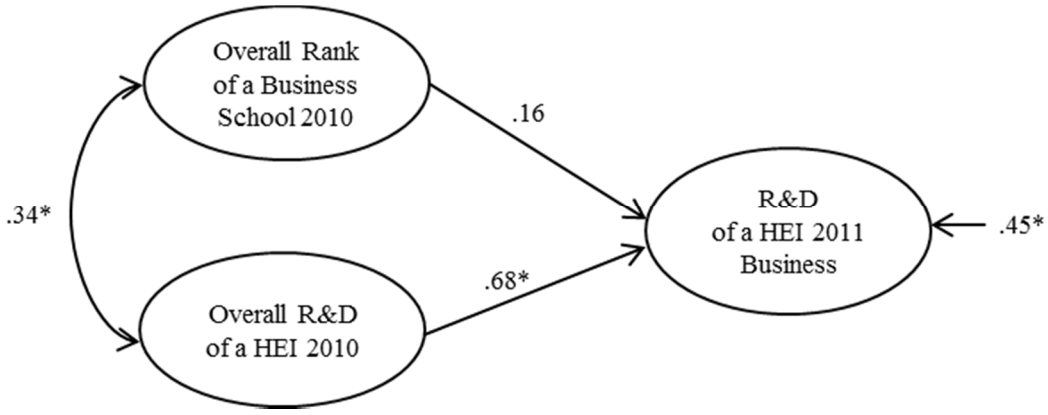


*Figure 15.* Completely standardized estimates for direct effects and endogenous disturbances for Structural Equation Model 4.

*$p < .05$.

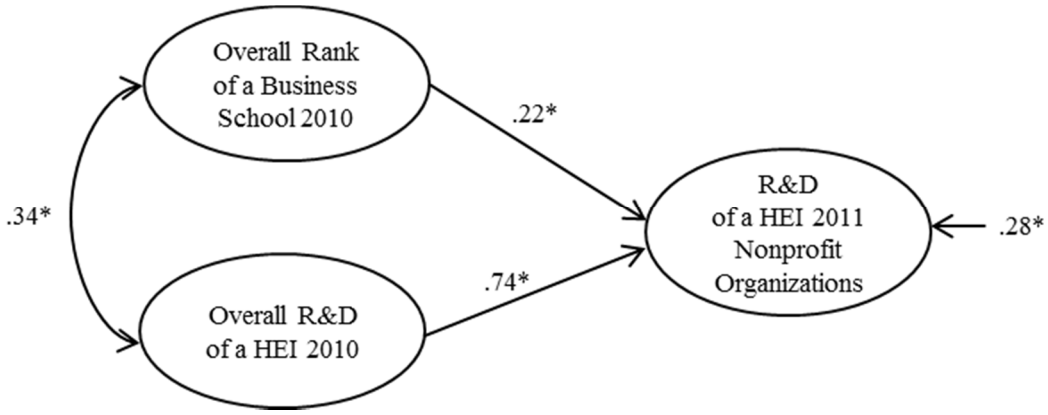


*Figure 16.* Completely standardized estimates for direct effects and endogenous disturbances for Structural Equation Model 5.

*$p < .05$.

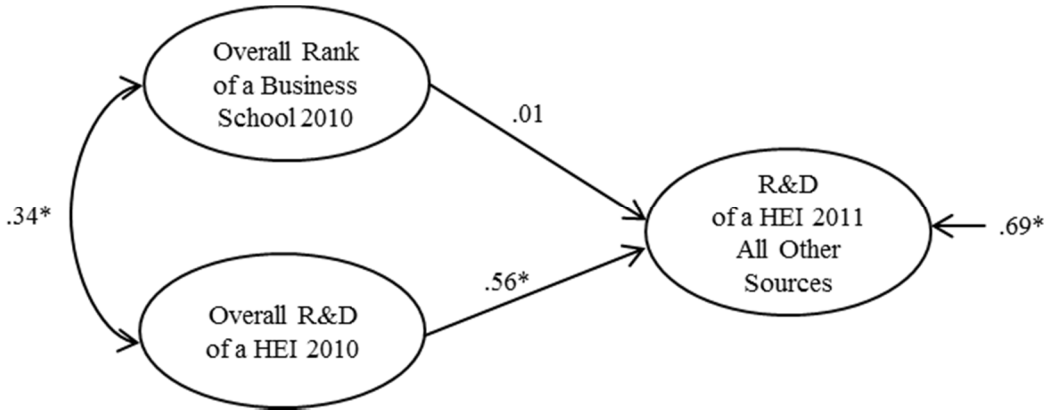


*Figure 17.* Completely standardized estimates for direct effects and endogenous disturbances for Structural Equation Model 6.

*$p < .05$.

for prior institutional resources (the overall R&D expenditure) in 2010, the rank of a business school in 2010 had a significant positive effect on R&D funding coming from

the federal government ($\beta$ = .15, $p$ < .05) and nonprofit organizations ($\beta$ = .22, $p$ < .05) in 2011. At the same time, Figures 13 and 14 demonstrate that the relationship between the rank of a business school and R&D funding coming from the state and local government and institutional funds was negative, with $\beta$ = -.40, $p$ < .05 and $\beta$ = -.41, $p$ < .05, respectively.

Contrary to the research hypothesis, the rank of a business school did not have any significant effect on either future R&D funding coming to a research university from businesses or other sources (see Figures 15 and 17). Institutional financial resources in a current year (the overall R&D expenditure of a higher education institution in 2010), however, had a significant positive effect, which ranged from moderate to large, on all the six future funding streams to a research university in a subsequent year (see Figures 12, 13, 14, 15, 16, and 17).

# CHAPTER V: DISCUSSION

The chapter puts the results of this study in the research and social contexts. The findings for each research question are discussed within the context of previous empirical and theoretical studies, and possible explanations, including social and economic forces, are provided for the phenomena under analysis. This chapter also delineates the unique contributions of this study to the higher education scholarship. Finally, this chapter specifies the study's limitations, articulates the ideas for future research on rankings, and suggests implications for practice for business school administrators, prospective M.B.A. students, and publishers of rankings of full-time M.B.A. programs. Based on the findings about reliability and validity of the business school rankings in this study, this chapter concludes that rankings of full-time M.B.A. programs can be viewed as a useful and credible tool to assist prospective M.B.A. students in their quest for a business program that best suits their needs.

## Research Question 1

**Ranking of rankings: Assessment of the full-time M.B.A. program rankings through the lens of the Berlin Principles on Ranking of Higher Education Institutions.** The lack of clear agreement in ranking of M.B.A. rankings produced by prospective students and higher education professionals may demonstrate different priorities and values of the two groups. However, placement of *The Financial Times* and the *Bloomberg Businessweek* M.B.A. rankings at the top of both students' and

professionals' lists (i.e., the first and second places, respectively, on both lists[34]) is probably not accidental. Both ranking systems include diverse indicators of business schools' quality, and this pluralistic approach to rankings methodology may be appealing to both prospective students and higher education professionals. *The Financial Times* global M.B.A. ranking system includes the largest number of quality criteria in its methodology among the five rankings under analysis. This ranking system creates the list of the 100 top M.B.A. programs in the world based on the analysis of three diverse areas that include alumni salaries and career development, the diversity and internationalization of the business school and its M.B.A. program, and research productivity of each school. The 20 criteria used to produce the school's rank include "weighted salary (US$)" and "salary percentage increase" (weighted 20% each), which is the information that appears to be very important for prospective students. The second most heavily weighted criterion (10%) is "the research rank", which is determined based on faculty research published in 45 academic and practitioner journals. This indicator

---

[34] Based on the raters' assessment of the five rankings, *The Financial Times* and the *Bloomberg Businessweek* M.B.A. rankings received 130.98 and 114.33 points, respectively, out of the maximum 180 points specified by the IREG methodology. Similarly, the results from the 2012 mba.com Prospective Students Survey (Graduate Management Admission Council, 2012b) indicated that 17% of prospective M.B.A. students considered *The Financial Times*, and 16% of prospective students regarded *Bloomberg Businessweek* to be extremely influential in their decision to apply to a particular M.B.A. program.

may be of special importance for higher education professionals. The remaining 50% of *The Financial Times* rankings are composed of value for money rank, careers progress rank, aims achieved rank, placement success rank, percentage of recent graduates employed at three months, alumni recommend rank, percentage of women faculty, percentage of women students, percentage of women on board, percentage of international faculty, percentage of international students, percentage of international members on board, international mobility rank, international experience rank, languages, percentage of faculty with doctorates, and *FT* doctoral rank. Thus, these indicators give a broad picture of the business school's performance, including students' outcomes, faculty quality, school's culture, government, and degree of internationalization. As a result, various indicators of an M.B.A. program's quality probably resonated with both prospective students' and higher education professionals' perceptions of a business school's quality.

The same argument is probably true about the *Bloomberg Businessweek* ranking system that is built based on students' and recruiters' surveys and faculty publications. Coupled with the facts that *Bloomberg Businessweek* produces the oldest M.B.A. rankings (debuted in 1988[35]), and it is one of the most influential business media

---

[35] As Lavelle (2012a) indicated, "*Bloomberg Businessweek* has been ranking the top U.S. full-time MBA programs every other year since 1988, and the top international programs since 2000" ("Best B-Schools Ranking History", para. 1). Nevertheless, *Bloomberg Businessweek* produces these rankings separately, and the rankings of the top international programs do not include U.S. full-time M.B.A. programs.

publications, it is not surprising that prospective students and higher education professionals have concurred with each other in their use and evaluation of the *Bloomberg Businessweek* rankings.

Prospective students' opinions of the *U.S. News & World Report* M.B.A. program rankings differ from those of higher education professionals, with students placing *U.S. News & World Report* second, and higher education professionals putting it fourth in their rankings. This finding about the popularity and influence of *U.S. News & World Report* on applicants is corroborated by the results of the study by Elliott and Soo (2013), who claimed that "US News rankings have a significant effect on applicants to US MBA programmes" (p. 162.), while it is not true for *The Financial Times* and *The Economist*. The explanation of the importance of *U.S. News & World Report* for prospective students in their decisions to apply to a particular program possibly lies in the fact that the *U.S. News & World report* rankings provide information about input-outcome measures (e.g., GPA, GMAT, school selectivity, salary, and employment rates), which allows applicants to benchmark their academic credentials and to project outcomes of their education. Higher education professionals, on the other hand, may be unwilling to accept the importance of the economic measures as outcomes of education. In this respect, in a qualitative study of 45 faculty and administrators of eight U.S. business schools, Dahlin-Brown (2005) unveiled a curious paradox of rejecting and accepting *U.S. News & World Report* as a proxy for academic excellence at the same time. As Dahlin-Brown (2005) noted, "While most participants agreed that the *U.S. News* ranking did not measure the

academic excellence of the schools they ranked..., they did feel that the rankings were generally correct in the schools they listed as the best" (pp. 171-172).

With respect to the *Forbes* rankings, one criterion – value added is probably too narrow to define the outcomes of education, which is admitted by both students and higher education professional in their rankings. However, students valued information provided in the *Forbes* ranking system more and placed it fourth in their rankings compared to the fifth position in the professionals' rankings.

The most striking differences in prospective students' and higher education professionals' views were observed in the evaluations of *The Economist* rankings. While higher education professionals placed *The Economist* rankings in the middle of their list (the third place), prospective students put it at the bottom of their rankings (the fifth place). Unlike *U.S. News & World Report* and *Forbes*, *The Economist* "concentrates on student expectations and experiences" (Bickerstaffe & Ridgers, 2007, p. 65). *The Economist* has a very elaborate methodology. The 13 rankings criteria are grouped into four categories, such as open new career opportunities, personal development/educational experience, increase in salary, and potential to network. The "open new career opportunities" category includes such measures as diversity of recruiters, assessment of career services, jobs found through the career services, and percentage of students in jobs three months after graduation. The "personal development/educational experience" category consists of four measures, such as faculty quality, student quality, student diversity, and educational experience. The "increase in salary" category consists of two measures, such as the increase in salary after graduation and leaving salary. Finally, the

"potential to network" category is comprised of such measures as breadth of alumni network, internationalism of alumni, and alumni effectiveness. Each of the 13 rankings criteria is assessed through corresponding indicators, which total 21. Educational experience, for example, is evaluated based on four indicators of quality, such as student rating of program content and range of electives, range of overseas exchange programs, number of languages on offer, and student assessment of facilities and other services. As a result, multiple indicators of each measure of quality allow a more comprehensive assessment of the quality of an M.B.A. program. It is not surprising that this complex methodology with a strong focus on the process of education, not the results, appealed to the raters from academia.

*The Economist* rankings lack factual information about M.B.A. programs, such as salary figures, for example. Moreover, it is not user-friendly, which is important for prospective students. The rankings webpage shows only the composite rank of M.B.A. programs. The criteria used to calculate the rank are not on the same page, which makes school comparison very difficult. As a result, *The Economist* was valued the least by M.B.A. applicants.

The differences in ranking of M.B.A. rankings produced by prospective students and higher education professionals bring into focus the lack of alignment of priorities and values between the two groups. Traditionally, academia relies on self-generated criteria of educational quality in the assessment of the quality of higher education institutions. Accreditation commissions consist of more experts from academe than the public. Therefore, their opinions about the standards of educational quality resonate with those of

higher education administrators and faculty. Perhaps, it is time to engage more stakeholders, including prospective students and alumni, recruiters, and experts from the corporate world, into the discussion of the quality of M.B.A. programs.

**Research Question 2**

**Reliability of the full-time M.B.A. rankings in the United States: Estimating predictability of changes of the rankings.** This study revisits the findings of the previous studies (Dichev, 1999; Fee et al., 2005) about the unpredictability of ranking systems of M.B.A. programs in the United States. The results of this study demonstrate that the *Bloomberg Businessweek,* the *Forbes, The Financial Times,* and the *U.S. News & World Report* rankings of the full-time M.B.A. programs are highly predictable and the future rankings are likely to demonstrate stability of either an upward or downward trend. First-order autocorrelations of error terms, which serve as "an empirical proxy for predictability" (Dichev, 1999, p. 203), were statistically significant and positive for all the four ranking systems under analysis. In the balanced data sets, first-order autocorrelations were $\rho = .80$, $p < .001$, $\rho = .77$, $p < .001$, $\rho = .70$, $p < .001$, and $\rho = .19$, $p = .012$ for the *Forbes*, the *U.S. News & World Report*, *The Financial Times*, and the *Bloomberg Businessweek* rankings, respectively. In the unbalanced data sets, first-order autocorrelations were smaller in magnitude, with $\rho = .76$, $p < .001$, $\rho = .73$, $p < .001$, $\rho = .57$, $p < .001$, and $\rho = .20$, $p = .001$ for the *Forbes*, the *U.S. News & World Report*, *The Financial Times*, and the *Bloomberg Businessweek* rankings, respectively.

The results of this study stand in total contrast to the results of the previous studies by Dichev (1999) and Fee et al. (2005) that provided evidence on the reversibility

of the *Businessweek* and the *U.S. News & World Report* rankings. Using the rankings data from the 1988-1996 and the 1990-2002 editions, both studies found that first-order autocorrelations were statistically significant and negative for both *Businessweek* and *U.S. News & World Report*, which indicated that rankings were not predictable.

The analysis of the predictability of rankings in this study yielded the following insights. First, stemming from the fact that previous studies used data from rankings editions before 2000, and this study used data from the 2000-2013 editions, it is possible to assume that methodological changes made to the four ranking systems between 2000 and 2013 have enhanced their reliability. Second, changes in rankings methodologies suggest that rankings are responsive to criticism and have made an effort to improve their methodologies.

In the absence of empirical research demonstrating that rakings have improved their methodologies, a brief excursus on the methodological changes in the *Bloomberg Businessweek* ranking system may provide evidence that rankings have attempted to improve their methodologies. In 1998, for example, the *Bloomberg Businessweek* rankings of full-time M.B.A. programs relied only on reputational assessment of business schools provided in the graduate and recruiter surveys and ranked only 25 top programs (*Bloomberg Businessweek*, 1998). In addition to a smaller scope of the inquiry into the metrics of the quality of M.B.A. programs in 1998, *Bloomberg Businessweek* surveyed a smaller number of M.B.A. graduates ($N$ = 9,598), who came from 61 top schools in the United States, and corporate recruiters ($N$ = 350), with a response rate of 63% ($n$ = 6,020) and 74% ($n$ = 259), respectively. The rankings were based on the three surveys of

M.B.A. graduates[36] and the three surveys of M.B.A. recruiters, which were conducted in 1998, 1996, and 1996. The graduate survey asked respondents 39 questions about the quality of teaching, curriculum, and placement. A weighted average of the three graduate surveys was calculated, with the 1998 survey contributing 50%, and the 1996 and 1994 polls contributing 25% each. The recruiter survey asked respondents to rank the top 20 schools at which they actively recruited, with 20 points given to school No.1. The total score for each school was divided by the number of responding companies that recruited M.B.A. graduates from that school[37]. In order to calculate the overall score for each school in 1998, *Bloomberg Businessweek* (1998) combined the raw scores from both surveys and gave a slightly greater weight to recruiters' opinion because there were "greater differences among the schools in the corporate survey" (*Bloomberg Businessweek*, 1998, "Crunching the Numbers: A Tale of Two Surveys", para. 6).

The methodology used by *Bloomberg Businessweek* to calculate its latest rankings of full-time M.B.A. programs in 2012 attests to its efforts to insure robustness of the

---

[36] Lavelle (2012b) provided the rationale for incorporating the results of the three most recent student surveys, claiming that it "effectively ensures that short-term issues – a new B-school building or unpopular dean – won't skew results one way or the other" ("How We Ranked the B-Schools", para. 4).

[37] Because larger schools tend to attract more recruiters, *Bloomberg Businessweek* (1998) "gave some of the smaller schools an upward adjustment to counter any possible bias" (*Bloomberg Businessweek*, 1998, "Crunching the Numbers: A Tale of Two Surveys", para. 5).

statistical analysis and to improve reliability of the rankings. First, following the major methodology change in 2000, the 2012 rankings were based on the three surveys of M.B.A. graduates, the three surveys of M.B.A. recruiters, which were conducted in 2012, 2010, and 2008[38], and a review of faculty research publications. The rank based on surveys of M.B.A. graduates and the rank based on surveys of M.B.A. recruiters contributed 45% each to the final ranking score. The rank based on the review of faculty research included publications in 20 top academic journals and faculty books reviewed in *The New York Times, The Wall Street Journal,* and *Bloomberg Businessweek* over a five-year period. The faculty research component contributed 10% to the final ranking, and the scores were adjusted for faculty size. As a result, *Bloomberg Businessweek* expanded the conceptual framework for the quality of M.B.A. programs by assessing quality through the lenses of the three main constituents of business schools: students, corporate recruiters, and faculty. Moreover, by assigning permanent weights to the three major components of its rankings, *Bloomberg Businessweek* removed an ambiguity in calculating the overall rank of M.B.A. programs in its previous rankings.

The second major difference in the 2012 *Bloomberg Businessweek* rankings methodology compared to its 1998 rankings is the magnitude of the ranking project

---

[38] Weighted averages of the three students and the three employer surveys were calculated, with the 2012 surveys contributing 50%, and the 2010 and 2008 surveys counting for 25% each (Lavelle, 2012b).

itself[39]. In 2012, *Bloomberg Businessweek* surveyed 18,640 M.B.A. graduates at 114 business schools in North America, Europe, and Asia, with a response rate of 56% ($n$ = 10,439). Combined with the student responses from 2010 and 2008, the total number of responses to the student survey amounted to 27,523, which constitutes almost twice as many responses as in 1998 ($n$ = 15,537). Moreover, the *Bloomberg Businessweek* rankings have become more democratic and inclusive, by moving from ranking only the top 25 elite business schools in 1998 to including 63 schools in the 2012 rankings. Thus, qualitative and quantitative methodological changes in the *Bloomberg Businessweek* rankings attest to the fact that rankings are evolving systems that make endeavors to better approximate the notion of business school quality.

The above-described methodological changes can also explain the findings of the present study about the predictability of rankings of full-time M.B.A. programs in the United States. While previous studies (Dichev, 1999; Fee et al., 2005) about the unpredictability of rankings of full-time M.B.A. programs used data from ranking systems produced before 2000, the present study utilized data from more recent rankings published between 2000 and 2013. As a result, the fact that rankings have been shown highly predictable in the most recent time period can be attributed to the changes in the psychometric properties of rankings of M.B.A. programs.

---

[39] It is interesting to note that, in 2012, as in 1998, *Bloomberg Businessweek* relied on expertise of the same statistical consultants, David M. Rindskopf and Alan L. Gross, professors of educational psychology at the City University of New York Graduate Center, to analyze the data.

Predictability of the rankings means that prospective M.B.A. students and corporate recruiters can rely on the rankings to assist in making their decisions. Nevertheless, this finding warrants a discussion of the connection between the predictability of rankings and the predictability of M.B.A. program quality. As institutions of higher education typically embrace change slowly (Rudolph, 1962; Torraco, 2005; Torraco & Hoover, 2005), it is tempting to make the same assumption about business schools: schools on a downward trend keep trending downward, and schools on an upward trend keep trending upward. As a result, predictability of rankings will serve as a signaling mechanism[40] of either an upward or downward trend of business school quality. Nevertheless, in response to intense competition from other business schools and alternative providers of business education and ever-increasing demands of the market, business schools are becoming more and more innovative in their educational strategies, including internationalization (Carlson, 2014; Hawawini, 2005); geographic expansion (Friga, Bettis, & Sullivan, 2003; INSEAD, 2010; Wharton, University of Pennsylvania, 2009); strategic alliances between schools (Carlson School of Management, University of Minnesota, 2010; Friga et al., 2003; Hay, 2008); education partnerships with corporations (Anderson, 2003); innovative classes (Stanley, 2010);

---

[40] Recent studies in marketing demonstrated the importance of brands and third-party information in signaling quality of the product in contemporary marketplaces. As Waldfogel and Chen (2006) noted, "Investments in brand, and third-party information, provide alternative methods for convincing potential customers that vendors will deliver as promised" (p. 425).

tailoring curricula to students' needs (de Onzoño & Carmona, 2007; Gloeckler, 2006) and the needs of different industries (Harvard Business School, 2010; Weidlich, 2001); interdisciplinary approach (de Onzoño & Carmona, 2007; Harvard will launch MD/MBA program, 2004); and changes in delivery format towards executive (de Onzoño & Carmona, 2007; Lorange, 2005), evening and part-time (de Onzoño & Carmona, 2007; Lorange, 2005), and online formats (Cornuel, 2005; Friga et al., 2003; Hawawini, 2005). Capitalizing on "the very favourable administrative and governance options offered to business schools (e.g. autonomy and separate naming)" by universities (Thomas, Lorange, & Sheth, 2013, p. 1992), some schools can make a striking change in academic standards and the learning process and improve their standing in the rankings. Alternatively, negative trends can affect the school quality, which will be reflected in the rankings.

A recent study by Gnolek, Falciano, and Kuncl (2014) made an interesting inquiry into the relationship between the *U.S. News & World Report* rankings and university quality by attempting to quantify "what amount of ranking change should be considered a meaningful difference, and what should be considered to be attributed to "noise" (p. 765). Gnolek et al. (2014) revealed that for universities ranked in the top 40 by *U.S. News & World Report*, fluctuations in rank within about ±2 should be considered noise. For universities ranked below the top 40, according to Gnolek et al. (2014), differences of ±4 (or even larger) should be attributed to noise. These findings demonstrated that "only when the rank moves beyond those levels can a statistically significant change be claimed" (Gnolek, Falciano, & Kuncl, 2014, p. 767). Thus, Gnolek, Falciano, and

Kuncl's (2014) research substantiated that "meaningful rank changes for top universities are difficult" (p. 778), and an immediate rise in rank may not indicate a real change in quality.

Business schools, being a constituent part of universities, are smaller in size, typically enjoy considerable operational and financial autonomy, and function in a competitive business environment. Therefore, a study of what constitutes a significant change in business schools quality versus noise in rankings of M.B.A. programs would equip business school administrators with meaningful information that can be used for strategic planning.

Meanwhile, straddling the fence in this respect, the implication for students planning to pursue an M.B.A. degree and their potential recruiters would be to approach rankings with caution. While in general the four rankings used in this analysis may be regarded as a signaling mechanism that shows predictable trends, the real dynamics of an M.B.A. program change is extremely intricate and depends on properties of multiple constituents of business schools, such as administrators, faculty, governance, alumni, businesses, etc.

**Reliability of the full-time M.B.A. rankings in the United States: Estimating reliability as stability of the ranking scores.** As reliability, which refers to "the accuracy or precision of a measurement procedure" (Thorndike & Thorndike-Christ, 2010, p. 118), is one of the most widely criticized flaws of the business school rankings methodologies (Dichev, 1999; Fee et al., 2005; Holbrook, 2007), this study has explored it from the two perspectives: reliability as predictability of changes of the rankings and

reliability as stability of the ranking scores. The results of the analysis of first-order autocorrelations of error terms showed that rankings are highly predictable and the future rankings are likely to demonstrate stability of either an upward or downward trend. In a similar fashion, the results of the correlational analysis of the five full-time M.B.A. ranking systems in the United States across 14 years strongly supported the hypothesis about the stability of the ranking scores, which also lends credence to the reliability of the rankings. Specifically, on average, for *Bloomberg Businessweek, The Economist, The Financial Times, Forbes,* and *U.S. News & World Report*, between 83% and 92% of the variance in the current year ranking was shared with the previous year's ranking. The results also indicated that levels and patterns of temporal stability in the five rankings of M.B.A. programs in the United States were comparable to those exhibited in prior studies (Collet & Vives, 2013; Devinney, Dowling, & Perm-Ajchariyawong, 2008; Iacobucci, 2013; Morgeson & Nahrgang, 2008). In particular, the magnitude of the correlation coefficients between ranking scores across 14 years for the *Bloomberg Businessweek* and the *U.S. News & World Report* rankings obtained in this study were identical to those found in the study by Iacobucci (2013), which corroborates the findings of the present research. With respect to *The Financial Times* rankings, a direct comparison between the magnitude of the correlation coefficients obtained in this study and those obtained by other researchers (Devinney et al., 2008; Iacobucci, 2013) was not possible due to the differences in the sample of business schools. While other studies examined the rankings of all business schools in *The Financial Times* Global M.B.A. rankings, this study focused only on the subset of the U.S. business schools in this ranking system.

Nevertheless, a general pattern of very high correlations (Iacobucci, 2013) and the shared variance between the current year ranking and the previous year's ranking (Devinney et al., 2008) had much in common with the results of this research.

The comparison of the results of the prior studies (Morgeson & Nahrgang, 2008) that examined the correlations for the *Bloomberg Businessweek* rankings between 1988 and 2004 and the current study that focused on the more recent editions of this ranking system between 2000 and 2012 allowed exploring the dynamics of changes in reliability of the *Bloomberg Businessweek* rankings over time. While Morgeson and Nahrgang's (2008) study revealed that the average correlation for *Businessweek* was .82 and, on average, 67% of the variance in the current year ranking was shared with the previous year's ranking, the results of this study demonstrated the weighted average correlation of .93 and 86% of the shared variance, on average. Therefore, there is empirical evidence to suggest that reliability of the *Bloomberg Businessweek* rankings has increased over time, which alludes to the fact that ranking systems are not stagnant and make an effort to improve their methodologies.

The specific contribution of this study is that it has broadened the research focus of the previous studies to include *The Economist* and the *Forbes* rankings into the analysis, which showed high levels of temporal stability similar to the *Bloomberg Businessweek*, *The Financial Times,* and the *U.S. News & World Report* rankings.

As Iacobucci (2013) observed, "The reliability results may be interpreted in a positive or negative manner" (p. 209). The positive attribute of a greater stability of the ranking scores is that it attests to good psychometric properties of the rankings.

Moreover, stability in the rankings over time benefits both prospective M.B.A. students and alumni who need stability in external validation of their degree over time and potential employers who rely on business school rankings in their recruitment decisions. By contrast, business schools may view stability of rankings as a major predicament because "reliability implies stickiness and difficulty in achieving enhanced placement in the rankings as a result of any efforts in program improvements" (Iacobucci, 2013, p. 209).

**Construct validity of the full-time M.B.A. rankings in the United States: Estimating construct validity with CFA of multitrait-multimethod matrices.** In terms of validity of the rankings, the results showed reasonable levels of evidence for convergent and discriminant validity of the *Bloomberg Businessweek* ranking system. These results were congruent with the findings of the recent study by Iacobucci (2013) about the validity of the *Bloomberg Businessweek, The Financial Times,* and the *U.S. News & World Report* rankings of M.B.A. programs. Although the methodologies employed to assess construct validity in this research and Iacobucci's (2013) study are different, their results seem to corroborate the idea of using the *Bloomberg Businessweek* rankings as a proxy for the quality of business schools.

In a few studies that have attempted to assess the psychometric properties of rankings of M.B.A. programs in the United States, researchers (Collet & Vives, 2013; Devinney et al., 2008; Morgeson & Nahrgang, 2008) focused primarily on the reliability aspects of the rankings. Iacobucci (2013) pioneered a new approach to the study of the quality of the rankings, which included "the use of multiple means of assessing both

reliability and validity" of the rankings (p. 204). Iacobucci (2013) examined convergent validity of the *Bloomberg Businessweek, The Financial Times,* and the *U.S. News & World Report* rankings by "seeking evidence that a measure should be correlated with variables that are related theoretically" (p. 211). Iacobucci (2013) theorized that seven constructs should be related to the quality of a business school: student input, student output, faculty research, reputation, undergraduate profile, institutional status, wealth, size, and university's greater locale. These seven latent constructs consisted of 29 indicators that measured the quality of a business school, including GMAT, GPA, M.B.A. graduates' salary, number of consulting jobs, number of finance jobs, number of nonprofit jobs, *Forbes* 5-year salary gain, Social Science Research Network (SSRN) (http://www.ssrn.com/en/) school rank, school all time cites (SSRN total number of school citations), school all time (SSRN total number of school papers), school authors (SSRN total number of school authors), measures of a school's reputation from two rankings: the *Businessweek* Undergraduate Business Schools Rankings and the *U.S. News & World Report* National Universities Rankings, a university's ACTs, SATs, SAT 25$^{th}$ percentile, SAT 75$^{th}$ percentile, undergraduate acceptance rate, 6-year graduation rate, a university's endowment, operating budget, tuition, private (vs. public), the size of the university as measured by its undergraduate and M.B.A. populations, school location in or near cities with *Fortune* 500 headquarters, population of a town/city, and the size of the greater metropolitan area where the school is located. Out of the 29 indicators of business schools' quality in Iacobucci's (2013) study, 18 indicators exhibited statistically significant correlations with the averaged rankings of M.B.A. programs from 2010 and

2012 in the *Businessweek* ranking system. The absolute values of these correlations ranged from .390 to .857. The correlations between the school rankings and indicators were both negative and positive, depending on the nature of the relationship with an indicator.

With regard to discriminant validity, Iacobucci (2013) posited that "the rankings should not be correlated with variables that are not theoretically related" (p. 214). These "extraneous variables" (Iacobucci, 2013, p. 214-215) included university ranks in sports (basketball, football, and soccer), locale resources (republican state and rent), university's weather (temperature range, July high temperatures, latitude, and longitude). As predicted, all these measures, except the university men's basketball team rankings, were not correlated with the rankings of M.B.A. programs in the *Businessweek* ranking system.

In contrast to Iacobucci's (2013) study, which used correlational analysis to examine convergent and discriminant validity of business school rankings, this research used a more complicated statistical procedure – multitrait-multioccasion confirmatory factor analysis that presents a variation of the traditional multitrait-multimethod matrices approach. Unlike Iacobucci's (2013) study that used arbitrary indicators of business schools' quality to measure construct validity of the *Bloomberg Businessweek* rankings, this research used indicators of business schools' quality specified by the *Bloomberg Businessweek* rankings methodology, which allowed avoiding ambiguity in defining the concept of business schools' quality. The specific attribute of this study also lies in the fact that its findings allowed validating *Bloomberg Businessweek's* methodology from a longitudinal perspective.

Despite existing flaws in their methodologies, it is evident that rankings of M.B.A. programs in the United States have evolved since their initial publications. One of the distinctive features of business school rankings is that they are produced by well-established business magazines, with the exception of *U.S. News & World Report*. Moreover, unlike most rankings of colleges and universities, business school rankings provide their users not only with rankings per se, but with a wealth of information published in magazines and online resources geared to the needs of business professionals. The media rankings, for example, offer featured profiles of selections of schools, blogs of successful M.B.A. students and graduates, interviews with business school constituents, such as deans, employers, financial experts, summaries of best practices as well as critiques of business schools, etc. Therefore, besides exploring the ranking scores, prospective M.B.A. students and recruiters should consult all available resources provided by the producers of rankings before making their decisions.

**Research Question 3**

**Conundrums of reputation: Anchoring effects in the full-time M.B.A. rankings in the United States.** Consistent with the research hypothesis, the results of the structural equation modeling demonstrated that the current university reputation had a strong impact on the future reputation of its business school. Moreover, consistent with the anchoring perspective, these associations solidified from year to year, $\beta = .89$, $p < .05$ in Model 1 and $\beta = .96$, $p < .05$ in Model 2. These results go in line with Cole and Lipton's (1977) and Safón's (2012) findings about the halo effect for schools of medicine and business affiliated with prestigious universities. The strong positive associations

between the university reputation and its business school reputation may be accounted for by the fact that reputation surveys of academic excellence in both the *U.S. News & World Report* rankings of national universities and M.B.A. programs are administered to top administrators: university presidents, provosts, and deans of admissions in the first case and business school deans and directors of accredited master's programs in business in the second case. The fact that both groups of administrators are guided by common university goals, strategic plans, and benchmarking results could introduce a group-thinking bias to business school administrators' judgment of academic excellence of other business schools and make their opinion aligned with university administrators. Nevertheless, it is necessary to mention that the reputational assessment of M.B.A. programs in the *U.S. News & World Report* rankings has a second component - the recruiter assessment score. Thus, the anchoring or halo effect impacts the opinion of both inside and outside stakeholders of business schools.

Although the current university reputation did not show a direct effect on the future business school rank in all three models, structural equation modeling allowed estimating the indirect and total effects. As demonstrated in Tables 25, 26, 27, the current university reputation had a strong positive indirect effect on the future business school rank through business school reputation (a mediator) in all three models, with $\beta = .84$, $p < .05$, $\beta = .87$, $p < .05$, and $\beta = .87$, $p < .05$, respectively. This finding is not surprising as reputation is one of the important components in the formulas for calculating both university and business school rankings. So, if the reputation of a business school is driven by the reputation of its parent university, and reputation

constitutes 40% of the business school rank, the rank of this business school will be indirectly affected by the reputation of the university. The absence of the direct effect of the reputation of a parent university on the rank of its business school may be explained by the fact that the business school rank includes other indicators of quality that constitute 60% of the total score.

The finding of the impact of the reputation of a parent university on the reputation of its business school resonated with the results of a qualitative study by Blackburn (2011) about factors that influence applicants' choice of an M.B.A. program. In his study, Blackburn (2011) revealed that prospective M.B.A. students value a university reputation more than a business school reputation when applying to an M.B.A. program. Moreover, Blackburn (2011) claimed that "there is an assumption that the standard of the business school can be inferred from the reputation of the university, that is, a 'good' university can be relied on to provide quality programmes in the individual schools" (p. 479). The findings of this study about a strong relationship between a parent university reputation and its business school reputation depict a two-faceted reality for business schools. On the one end of the spectrum, business schools affiliated with prestigious universities enjoy the halo effect irrespective of their actual reputation. On the other side of the spectrum, the fact that the reputation of a university is a major contributing factor to a prospective M.B.A. student's choice of a business school, and it is not easy to change a business school reputation because it is anchored to its parent university reputation (Safón, 2012) poses a significant challenge to business school administrators in terms of strategies to enhance quality and reputation.

**Research Question 4**

**Are we making money for the entire institution? The impact of rankings of full-time M.B.A. programs on funding streams to institutions of higher education.** While Research Question 3 in this study delved into the problem of the halo effect of the reputation of a parent university on the reputation of its business school, this part of the study changed the vector of inquiry and explored the "peripheral halo effect" (Pitsakis, Souitaris, & Nicolaou, 2015, p. 321): how rankings of full-time M.B.A. programs affect funding streams to institutions of higher education. Although "business schools are often regarded as "cash cows" at American universities, because they offer an important revenue for many institutions, often operating at 10-20% positive operating margin" (Röbken, 2004, p. 33), their impact on resource flows to universities other than tuition and fundraising remains entirely unexplored. As "reputation is clearly an organizational resource in higher education" (Bastedo & Bowman, 2011, p. 19), a higher standing in rankings of full-time M.B.A. programs was predictive of greater R&D funding coming to universities from the federal government ($\beta = .15$, $p < .05$). While the connection between rankings of universities and federal funding for R&D was supported by empirical study (Bastedo & Bowman, 2011), the link of the rankings of business schools and federal money for R&D coming to their parent institutions is not easy to explain. On the one hand, this positive effect may be explained by the money that business schools bring to their parent universities through government grants. On the other hand, the high correlation between the reputation of business schools and the reputation of their parent universities makes it difficult to disentangle the causes of this effect. In any case, the

reality is that, in 2011, for example, approximately 20% of the NSF's spending on grants for R&D went to just 10 highly ranked top research universities (Weigley & Hess, 2013).

While "The federal government doles out money through earmarks, which are related to political power in legislative districts, and through government agencies, which most often employ experts to dole out funds on a project-by-project basis" (Bastedo & Bowman, 2011, p. 20), state funding presents a more complicated phenomenon that is shaped by a wide range of economic, political, cultural, and institutional factors, and even rankings (Institute for Higher Education Policy, 2009; Weerts & Ronca, 2006; Weerts & Ronca, 2012; Weerts, Sanford, & Reinert, 2012 ). Whatever the amalgam of the local forces that determine fiscal policies in various states, the results of this study demonstrated a negative relationship between rankings of business schools and the resource flow for R&D to their parent universities from state and local governments ($\beta$ = -.40, $p$ < .05). Affluent higher education institutions may be less likely to obtain funds for R&D from the state under the pressure of other fiscal priorities.

While it is tempting to explain this result by the fact that 26 out of the 50 institutions in the sample are private institutions and cannot count on state appropriations as a major source of funding, some of the private institutions receive generous funding for R&D from the state and local government. In 2011, for example, Duke University received $31,519,000 for R&D from the state of North Carolina, while the University of North Carolina at Chapel Hill received only $6,477,000 (National Science Foundation/National Center for Science and Engineering Statistics, 2013). In the same year, Stanford University received $39,148,000 from the state of California, which is

only 1.5 less money than the University of California, Berkeley ($59,621,000) (National Science Foundation/National Center for Science and Engineering Statistics, 2013).

An interesting finding in this study showed a moderate negative effect of rankings of business schools on the R&D money that comes from institutional funds ($\beta$ = -.41, $p < .05$). It seems that richer universities are reluctant to invest their own money in the R&D projects. Statistics demonstrate that institutions with massive endowments and lucrative federal funding rely mostly on government funds for R&D. In 2011, for example, 88% of the R&D money at John Hopkins University came from federal funding. Similarly, at the University of Pennsylvania, 80% of the R&D budget came from federal funds (Weigley & Hess, 2013).

Another interesting finding was that nonprofit organizations appeared to be susceptible to standing of business schools in the rankings ($\beta$ = .22, $p < .05$) when it comes to R&D funding of higher education institutions, while businesses were impervious to rankings pressure ($\beta$ = .16, $p > .05$). One possible explanation of this finding lies in the fact that there has been an increase in the number of business schools and schools of business and public administration offering programs in nonprofit management (Mirabella, 2007), and their graduates working in the nonprofit sector are more likely to be affected by M.B.A. rankings and university prestige when allocating money for university R&D projects. Businesses, on the other hand, are guided by a number of considerations, including profit, intellectual ownership, partnership responsibilities, liabilities, etc., the complicated interplay of which may or may not include rankings when investing money in university R&D projects.

**Limitations**

A number of limitations in this study warrant mention. The most notable limitation pertains to the quality of the data utilized in the study. First, the author of the study used publicly available rankings data of full-time M.B.A. programs in the United States from the five ranking systems: *Bloomberg Businessweek*, *Forbes*, *The Economist*, *The Financial Times*, and *U.S. News & World Report*. As a result, the author did not have any control of the data collection, processing, and validation in the ranking systems. Moreover, the five ranking systems of M.B.A. programs do not provide any information about the magnitude of the measurement error or reliability of their ranking scores which may complicate the results of the study.

Second, inadequate sample size, sometimes as small as 42 observations, brought several methodological problems. The first methodological problem is method of estimation. For research question two that examined construct validity of the *Bloomberg Businessweek* ranking system with CFA of multitrait-multimethod matrices and used rank-ordered data, nonparametric methods, such as robust maximum likelihood, weighted least squares, or elliptical reweighted least squares, would be more appropriate than maximum likelihood. However, because these methods of estimation require a large sample size (Boomsma & Hoogland, 2001), estimation errors occurred when the author tried to apply them.

The second methodological problem caused by the small sample size was the limited number of variables that could be used in the proposed models. Tanaka (1987) asserted that the appropriateness of sample size should be linked to the size of the model

to be estimated, more specifically, "to the number of parameters estimated in the model" (p. 143). As the ratio of sample size to the number of free parameters estimated should meet the minimum requirement of 5 to 1, as recommended by Bentler and Chou (1987), which can be even more conservative in the opinion of other researchers depending on the research design (MacCallum et al., 2001), the small sample sizes in the rankings prevented the author from constructing more comprehensive models for research questions three and four. Initially, research question three that examined the anchoring effect of rankings of higher education institutions and their reputation on rankings of their business schools and their reputational assessment had a more complex conceptual framework. The original conceptual framework included two more latent variables: the reputational assessment of business schools at Time1 (the exogenous variable) and the overall rank of business schools at Time 1 (the endogenous variable). The reputational assessment of business schools was represented by a latent construct that was operationalized as the peer assessment score and the recruiter assessment score in a particular year. Overall, the inclusion of these variables and the operationalization of the latent variables as represented by the two observed variables would add 13 free parameters to be estimated in each of the three models. The total number of free parameters to be estimated in each of the three original models was 27. As the sample size totaled 64 business schools, the minimum requirement for the ratio of sample size to the number of free parameters estimated was not met. Model simplification resulted in 14 free parameters, which made parameter estimation possible.

In a similar vein, the original conceptual framework for research question four that explored how rankings of full-time M.B.A. programs affect funding streams to institutions of higher education was represented by one model that included nine variables: the overall rank of a business school in 2010, type of institutional control, the overall R&D expenditure of a higher education institution in 2010, R&D funding from the federal government, R&D funding from the state and local government, R&D funding from institutional funds, R&D funding from business, R&D funding from nonprofit organizations, and R&D funding from all other sources in 2011. The overall rank of a business school was viewed as a latent construct as measured by the *U.S. News & World Report*, *Bloomberg Businessweek*, *The Financial Times*, *The Economist*, and the *Forbes* rankings. R&D funding from the federal government, R&D funding from the state and local government, R&D funding from institutional funds, R&D funding from business, R&D funding from nonprofit organizations, and R&D funding from all other sources in 2011 were the endogenous variables. Overall, the structural equation model had 43 free parameters to be estimated. This overarching model was simplified by removing one variable (type of institutional control) and operationalizing the overall rank of a business school as measured by the *U.S. News & World Report* rankings only because the number of business schools that appeared in all the five rankings was small. Keeping the same six endogenous variables, the overarching model was divided into six models, and each of the models had eight free parameters to be estimated. As the total sample size equaled 51 business schools, the minimum requirement for the ratio of sample size to the number of free parameters estimated was met. Thus, the methodological problem, rather than

flaws in the conceptual framework, accounted for rather simple path analysis models that explored anchoring effects in the full-time M.B.A. rankings and the impact of rankings of full-time M.B.A. programs on funding streams to institutions of higher education.

Besides the small sample size, another formidable limitation – type of the rankings data – was present in the study. Only two ranking systems, *Bloomberg Businessweek* and *U.S. News & World Report*, provided both ranks (rank-ordered data) and composite scores (continuous data) for M.B.A. programs. As Kline (2011) specified that the endogenous variables be continuous in structural equation modeling, this requirement limited the choice of ranking systems that could be used in the models that examined anchoring effects in the rankings of M.B.A. programs in research question three.

Finally, data availability presented a challenge in this study. Some ranking systems, *The Economist* and *U.S. News & World Report*, for example, while updating information about rankings on their websites every year, block access to data from previous years. While these M.B.A. rankings also publish rankings in their printed magazines, the printed versions are curtailed in content compared to their online versions. For example, *U.S. News & World Report* provides information about more than 100 business schools online, but includes around 50-70 business schools in its paper publication of rankings. This number of schools was even smaller in *U.S. News & World Report* magazines published before 2000. Thus, longitudinal analysis presented a certain challenge in this study.

The availability of financial data was also a major hurdle for research question four in this study. As this study was not sponsored, the author relied on publicly available financial data that restricted the choice of indicators in the model examining the impact of rankings of full-time M.B.A. programs on funding streams to institutions of higher education. Research and development funding (R&D) data, which included R&D funding coming from the federal government, the state and local government, institution funds, business, nonprofit organizations, and from all other sources, were collected from the National Science Foundation's Higher Education Research and Development Survey 2011 (National Science Foundation/National Center for Science and Engineering Statistics, 2013). Other potential financial indicators could include institutions' funding from foundations, corporations, donations from alumni, fundraising, etc. This information on private giving is collected by the Council for Aid to Education (http://cae.org/), but its data access is rather expensive. As a result, data permitting, the ability to use all five ranking systems of M.B.A. programs and a wider range of financial indicators would yield a more comprehensive understanding of the influence of business schools on resource flows to their parent universities as well as a more in-depth analysis of the psychometric properties of rankings.

Although Krippendorff's α met the minimum acceptable level of the agreement estimate of interrater reliability of .667 for "tentative conclusions" (Krippendorff, 2004a, p. 241) for all the ranking systems except *Forbes* (α = .666), the low interrater reliability estimates (.681 was the highest estimate) can be considered as a limitation in research question one. Hensel, Meijers, van der Leeden, and Kessels (2010) demonstrated that the

reliability level of ratings increased with "an increasing number of raters" (p. 2818). Hensel et al. (2010) also showed that the strongest increase in reliability was observed between two and six raters. In case when more than six raters were involved, Hensel et al. (2010) noted that the increase in reliability diminished. Therefore, the increase in the number of raters from three to six would enhance the agreement estimates of interrater reliability in this study.

**Contributions of the Study**

While the results of this study support previous findings, this research makes several unique contributions to higher education scholarship. First, the present study revealed a discord in students' and higher education professionals' views of the M.B.A. program quality which is operationalized by the five rankings of M.B.A. programs. While students prize input-output information and opinions of alumni and recruiters about the program value in the rankings, higher education professionals focus on various aspects of the educational process, such as teaching quality, etc., which aligns with the traditional perception of quality in academia and accreditation provisions. This finding should encourage business school leaders to engage in dialogue about emerging needs of prospective M.B.A. students. It also shows the necessity to calibrate a complicated system of business education involving complex and often conflicting priorities of its stakeholders.

Second, the specific attribute of this study consists in evaluating the psychometric properties of the five full-time M.B.A. rankings in the United States as the emerging sources of influence on the public opinion with regard to the quality of business schools.

The results of the study demonstrate that the M.B.A. ranking systems are highly predictable and the future rankings will demonstrate stability of either an upward or downward trend, which revisits the traditional perspective on the M.B.A. rankings (Dichev, 1999; Fee et al., 2005) as unpredictable systems. Predictability of rankings, coupled with findings of this study about longitudinal stability of ranking scores, adds credence to reliability of the rankings. The findings about reliability of rankings may testify to the positive dynamics of changes in rankings methodologies, which shows that rankings are responsive to criticism and have made an effort to improve their methodologies.

Third, one of the most unique aspects of this study is the confirmation of the notion of construct validity of the *Bloomberg Businessweek* ranking system as a proxy for the quality of business schools. Moreover, the notable contribution of this research lies in the fact that it allows validating *Bloomberg Businessweek's* methodology from a longitudinal perspective. Methodologically, this research demonstrates how higher education researchers can apply complicated statistical procedures, such as confirmatory factor analysis of multitrait-multimethod matrices, which is traditionally used in other areas of social sciences: psychology (Biesanz & West, 2004; Langer et al., 2010) and organizational research (Bagozzi & Phillips, 1982; Bagozzi et al., & 1991), to the study of complex organizational issues in higher education.

Fourth, this study fills a gap in the limited research literature on the anchoring effect in the reputation of professional schools, especially schools of business. The results of the present research demonstrate that the current university reputation has a

strong impact on the future reputation of its business school. Moreover, in consonance with the anchoring perspective, these associations solidified from year to year. In this vein, the findings of this study may be of interest to both major stakeholders within the system of business education, business schools and their students whose future careers may be affected by the prestige of their degree-granting school as impacted by the ranking of its parent institution.

Finally, although research on various financial benefits associated with high standing in both the university and professional school rankings has carved out a special niche in the higher education scholarship, the problem of the impact of rankings of full-time M.B.A. programs on funding streams to institutions of higher education has remained an obvious lacuna in empirical research. In this light, this is the first study that explored the "peripheral halo effect" (Pitsakis et al., 2015, p. 321) that is how a part of the institution (a full-time M.B.A. program) affects funding streams to the entire institution of higher education. Disentangling the differential impact of the rank of a business school on R&D funding coming to the university from the federal, state and local governments, institution funds, business, nonprofit organizations, and other sources may be considered a useful contribution to the discussion of the value of rankings in both research literature and institutional practice.

**Implications for Future Research on Rankings**

The limitations of this study and the gap in research literature on the psychometric properties of rankings open a number of potential avenues for future research. First, while this study found that the *Bloomberg Businessweek,* the *Forbes, The Economist, The*

*Financial Times,* and the *U.S. News & World Report* M.B.A. rankings are highly stable over time, it might be interesting to explore the levels of temporal stability for different quartiles of these ranking systems. In this respect, by calculating the percentage of schools that remained in the same quartile from one year to another from 1999 until 2010, Collet and Vives (2013) found "high levels of stability within 1st quartile" (p. 546) for both U.S. (91.50%) and non-U.S. (84.40%) schools in *The Financial Times* M.B.A. rankings. The degree of stability decreased with the quartile increase for both groups of business schools. Stemming from Collet and Vives' (2013) findings and the results of the current study, it might be interesting to explore the differences in the levels of longitudinal stability based on the quartiles in the *Bloomberg Businessweek,* the *Forbes, The Economist, The Financial Times,* and the *U.S. News & World Report* M.B.A. rankings, which are stable overall. The differential analysis of the five rankings based on quartiles and grounded in a different methodology than Collet and Vives' (2013) study – the use of Spearman rank correlations – might provide a better understanding of rankings as "stable hierarchies" (Collet and Vives, 2013, p. 546).

Second, it might be interesting to research how the rank of a business school impacts the rank of its parent university and corroborate Safón's (2012) conclusion that "the reputations of schools and their parent universities are influenced reciprocally" (p. 177). In this respect, Pitsakis, Souitaris, and Nicolaou (2015) introduced the concept of the "peripheral halo effect" (p. 321) to show how a part of the university – academic spinoffs – can "generate income for universities not only directly via equity positions but also indirectly via reputational benefits" (p. 321). Therefore, the peripheral halo effect of

business schools might be an interesting phenomenon which allows delving into the complex sphere of intraorganizational relationships.

Third, while this research demonstrates the anchoring effect in the full-time M.B.A. rankings, it might be interesting to explore the anchoring effect in other professional fields, for example, law or medical schools. Researchers might broaden the hypothesis to determine particular fields where there is a stronger relationship between parent university rankings and their professional school rankings. Moreover, the set of models for exploring the anchoring effect is not exhaustive, and there are other ways of asking the research questions in this study that require different data collection and analytic approaches. In this respect, increasing the number of lags in the model of the anchoring effect in the full-time M.B.A. rankings might be a way to improve the model because universities embrace change slowly, and lag 2 might not represent enough time to reveal the impact of this change.

Fourth, it might be interesting to quantify "noise" and substantiate change in school quality in the five ranking systems of M.B.A. programs. In this vein, Gnolek et al. (2014) determined that for universities in the top 40 in the *U.S. News & World Report* college rankings, differences in rank within ± 2 should be considered noise. Differences of ± 4 (or even greater) should be considered noise for universities ranked beyond 40. Stemming from these findings, future research might examine the amount of change that would be required for a business school to improve its rank or move into the top 20 in the *Bloomberg Businessweek,* the *Forbes, The Economist, The Financial Times,* and the *U.S.*

*News & World Report* M.B.A. rankings, which would suggest a real improvement in quality beyond "noisy" fluctuations in rankings.

Fifth, the small sample size prevented the author of the study from examining the relationship between rankings of business schools and funding streams to their parent universities for public and private institutions separately. Therefore, as rankings include more institutions in their lists, it might be interesting to explore this relationship within the context of public and private education, which would lead to a more conceptually rich understanding of the association between reputation and funding in higher education.

Six, some methodological approaches used in this research to explore business school rankings may be applied to the study of rankings of other professional schools. While this study attempted to unravel the differential impact of a rank of a business school on R&D funding coming to its parent university from the federal, state and local governments, institution funds, business, nonprofit organizations, and other sources, it might be interesting to explore this impact in the context of medical school rankings.

Seven, previous research showed that a rise in the rankings allows business schools to increase tuition, attract high quality applicants, garner resources from alumni and corporate donors, and attract prestigious companies to recruit their graduates (Argenti, 2000; Corley & Gioia, 2000; Eberhardt & Moser, 1997; Rindova et al., 2005; Safón, 2007). Moreover, Fee et al. (2005) revealed that business school deans turnover increases when a business school drops in the rankings. Dahlin-Brown (2005) provided evidence that rankings have "a distinct impact on the size and staff of the career services and admissions offices" (p. 169). Although the impact of rankings on various aspects of

business school operations is well-documented, there is lack of research on how rankings of business schools affect institutional decision making. In this light, possible avenues for future research might include the study of institutional efforts to improve the quality of their business schools, such as additional funding, branding, etc.

Finally, scholars must explore alternative approaches to analyze and understand the performance of business schools. Goldstein, Lunn, and Peng (2015) admitted that "Comparative analyses are powerful tools to understand the differences between institutions, but they are challenging to carry out" (p. 603). Rankings seem to have successfully carved out a niche for themselves in this field. As a result, the analysis performed by the *U.S. News & World Report* rankings has become, for example, "the default to compare U.S. medical schools" (Goldstein, Lunn, & Peng, 2015, p. 603). Therefore, scholars should suggest other models of comparative analysis of business schools that would unveil changes in the M.B.A. market place and inform the public about the quality of M.B.A. education. In this respect, one of the possible approaches might be the analysis of the M.B.A. rankings correlations with other measures of business school quality.

**Implications for Practice**

Contrary to a popular sentiment against rankings in higher education, institutional leaders should realize "the role of the popular press in the legitimization of US business schools" (Drnevich, Armstrong, Crook, & Crook, 2011, p. 183). As Sweitzer and Volkwein (2009) admitted, "even if institutions follow the will of many in the higher education community by deemphasizing their ranking, the consumer's fascination with

finding out "Who's Number One?" may be the deciding force in determining whether or not rankings at any level are here to stay" (p. 834). In this light, the findings of this research have important practical implications for business school leaders when setting strategic goals and communicating expectations, prospective M.B.A. students when making their application decisions, and publishers of rankings when making changes to rankings methodologies.

A practical recommendation for business schools would be to collaborate with publishers of rankings and business professionals to improve the psychometric properties of rankings. As Dichev (2008) noted, "If schools are in charge of designing the rankings, they have every incentive to do it right, and there can be no complaints later that rankings are not fair" (p. 224). In this respect, as suggested by one of the business professionals, one of the changes to the ranking systems of business schools would be to include "a criterion that reflects practical relevance, vitality and impact. A new approach to measurement and metrics should reflect society's complex expectations of business schools" (Denning, 2014, What to do? para. 4).

The finding of the impact of the reputation of a parent university on the reputation of its business school is congruent with Safón's (2012) research and brings up two thought-provoking implications. First, the fact that a business school reputation is anchored to its parent university reputation raises the question of "the efficiency of short-term strategies" (Safón, 2012, p. 169) aimed at increasing reputation and standing in the rankings without improving quality. Second, because of the reciprocal relationship between the reputations of business schools and their parent universities, as shown by

Safón (2012), it seems sensible to join forces to exploit the obvious synergies between the two parties by "modifying, where possible, their communication strategies towards the outside world with a view to associating the two names" (Safón, 2012, p. 177). For instance, while providing very detailed information about schools' standing in various business school rankings (e.g., rankings of best undergraduate business degrees, rankings of full-time, part-time, and executive M.B.A. programs, rankings of M.B.A. specialties, such as accounting, finance, marketing, entrepreneurship, information systems, etc.), websites of business schools do not offer any information about the overall rankings of their parent universities. Nevertheless, it is evident from research that business schools can benefit from the halo effect (Blackburn, 2011; Cole & Lipton, 1977; Safón, 2012) because prospective M.B.A. students often assume that "the standard of the business school can be inferred from the reputation of the university" (Blackburn, 2011, p.479).

The results of the analysis of psychometric properties of rankings of full-time M.B.A. programs in this research, such as high predictability and stability of either an upward or downward trend in future rankings, high temporal stability, and reasonable levels of evidence for convergent and discriminant validity, have an important implication for prospective students that counters a strong sentiment against rankings in previous research literature. Taking into account both strengths and weaknesses of each of the five ranking systems of M.B.A. programs in the United States, prospective students can portray a comprehensive picture of the business school "market" and personalize a search for a school that meets their needs. As Glenn Hubbard (2015), Dean of Columbia Business School, put it, "finding the top business school for you doesn't have to be a

guessing game. The information needed to identify the best schools is out there, and if you ask this dean, that's the real business of business school rankings" (Do b-school rankings really matter? para. 10).

In summary, it is challenging to come up with a definitive recommendation for the users of rankings with respect to what ranking system they should prioritize in their decision making because this research did not explore all the aspects of reliability and validity in all the five ranking systems. For instance, this study examined convergent and discriminant validity of the *Bloomberg Businessweek* full-time M.B.A. rankings because only this ranking system met the methodological criteria of CFA of multitrait-multioccasion matrices. On reflection, however, my judgment is that *The Financial Times* full-time M.B.A. ranking system is the strongest in the set of the five rankings analyzed in this study due to several reasons. First, both prospective M.B.A. students and higher education professionals agreed about the value of *The Financial Times* rankings by placing *The Financial Times* at the top in the ranking of rankings. Second, *The Financial Times* demonstrated predictability of changes, with significant positive first-order autocorrelation in both balanced $\rho = .70$, $p < .001$ and unbalanced $\rho = .57$, $p < .001$ data sets. Third, *The Financial Times* showed stability of the ranking scores, with the weighted average Spearman rank order correlation of .91. Being the two facets of reliability, predictability of changes and stability of scores testified to the good psychometric properties of *The Financial Times* ranking system. Fourth, *The Financial Times* global M.B.A. ranking system includes the largest number of quality criteria in its methodology among the five ranking systems in this study and provides a broad picture

of the attributes of business schools. These criteria consist of both objective (e.g., weighted salary) and subjective (e.g., aims achieved rank) criteria of business school quality, without overemphasizing either objective metrics of the school quality or subjective judgment of the respondents to ranking surveys. Finally, by creating the list of the 100 top M.B.A. programs in the world, *The Financial Times* rankings expand M.B.A. applicants' horizons and provide insight into the quality of business education worldwide.

**Conclusion**

With an increased attention to the economic success of college graduates due to the unprecedented levels of student indebtedness and high unemployment rates, government agencies started to publicize performance of higher education institutions on accountability measures, such as graduation rates, percentage of graduates employed, and their salaries. In Minnesota, for example, the Minnesota Department of Employment and Economic Development and the Minnesota Office of Higher Education launched a new online tool - the Graduate Employment Outcomes (http://mn.gov/deed/data/data-tools/graduate-employment-outcomes/) that "provides a wealth of information on how recent graduates in different majors fared in the Minnesota job market" (Leibert, 2014, Introducing the Graduate Employment Outcomes tool, para. 2). The Graduate Employment Outcomes website offers, for example, information about the percentage of M.B.A. graduates employed and their average earnings during the second, third, and fourth year from graduation from schools of business in Minnesota. The website provides even more detailed information about educational outcomes, such as the number of graduates, their median hourly wages, top industry of employment, and top region of

employment one, two, and four years after graduation for every field (e.g., business administration, management, and operations; marketing; taxation; insurance; accounting and related services; finance and financial management services; etc.) in business education at every university and college in the state of Minnesota. From this perspective, it is necessary to admit that rankings of full-time M.B.A. programs in the United States provide more comprehensive information about the economic success of M.B.A. graduates, for example, return on investment, time necessary to pay back, employment rate at graduation, tuition, base salary by occupation, industry, geographic region, citizenship status, etc., than government agencies. Moreover, rankings provide information about the economic success of M.B.A. graduates tailored to each particular business school, and this information is nationally available. Rankings have also become more inclusive, currently providing available information not only about top but also regional business schools[41]. Additionally, online versions of business school rankings offered by the media are a user-friendly benchmarking tool that allows prospective students to compare M.B.A. programs based not only on the rank score but also on other criteria. Therefore, rankings of full-time M.B.A. programs can be viewed as a useful and

---

[41] For the 2016 Best Business Schools rankings, *U.S. News & World Report* surveyed all 464 master's programs in business accredited by the Association to Advance Collegiate Schools of Business International. A total of 385 schools responded, and only 126 schools provided enough of the data needed to calculate the full-time M.B.A. rankings. Nevertheless, all 464 schools appear in the online directory of *U.S. News & World Report* (Flanigan & Morse, 2015).

credible resource to assist prospective M.B.A. students in their quest for a business program that best suits their needs.

In the contemporary landscape of business education, it is evident that rankings are becoming increasingly visible and important in both national and international contexts. It is reasonable to see changes in the M.B.A. program rankings, and these changes are probably more than the noise in the system. Although rankings of M.B.A. programs have honed their unique niche as a measure of quality of business education, the fundamental question of whether rankings really reflect what their users think – the quality of business education – still remains a conundrum, and the challenge for the producers of rankings is to do it correctly.

# REFERENCES

AACBS International. The Association to Advance Collegiate Schools of Business. (2005). *The business school ranking dilemma.* Retrieved from http://www.gfme.org/issues/pdfs/Media%20Rankings.pdf

Ahn, T. S. (1987). *Efficiency and related issues in higher education* (Unpublished doctoral dissertation). University of Texas, Austin, TX.

Ahn, T., & Seiford, L. M. (1993). Sensitivity of DEA to models and variable sets in a hypothesis testing setting: The efficiency of university operations. In Y. Ijiri (Ed.), *Creative and innovative approaches to the sciences of management* (pp. 191-208). Westport, CT: Quorum Books.

Altbach, P. (2006). The dilemmas of ranking. *International Higher Education, 42*, 2-3. Retrieved from https://htmldbprod.bc.edu/prd/f?p=2290:4:0::NO:RP,4:P0_CONTENT_ID:100222

Althauser, R. P., & Heberlein, T. A. (1970). Validity and the multitrait-multimethod matrix. *Sociological Methodology, 2*, 151-169. Retrieved from http://www.jstor.org/stable/pdfplus/270788.pdf

Alwin, D. F. (1974). An analytic comparison of four approaches to the interpretation of relationships in the multitrait-multimethod matrix. *Sociological Methodology, 5*(1973-1974), 79-105. Retrieved from http://www.jstor.org/stable/pdfplus/270833.pdf

Anderson, L. (2003, March 24). Collaboration, not rivalry, is the best way ahead: Corporate universities vs business schools. *Financial Times,* p. 4.

Argenti, P. (2000). Branding B-schools: Reputation management for MBA programs. *Corporate Reputation Review, 3*(2), 171-178. doi:10.1057/palgrave.crr.1540111

Asteriou, D., & Hall, S. G. (2011). *Applied econometrics* (2nd ed.). New York, NY: Palgrave Macmillan.

Astin, A. W. (1985). *Achieving educational excellence*. San Francisco, CA: Jossey-Bass.

Badenhausen, K. (2011, August 3). *Business school ranking methodology.* Retrieved from http://www.forbes.com/sites/kurtbadenhausen/2011/08/03/business-school-ranking-methodology/

Bagozzi, R. P., & Phillips, L. W. (1982). Representing and testing organizational theories: A holistic construal. *Administrative Science Quarterly, 27*(3), 459-489. Retrieved from http://www.jstor.org/stable/2392322

Bagozzi, R. P., Yi, Y., & Phillips, L. W. (1991). Assessing construct validity in organizational research. *Administrative Science Quarterly, 36*(3), 421-458. Retrieved from http://www.jstor.org/stable/2393203

Bailey, J., & Ford, C. (1996). Management as science versus management as practice in postgraduate business education. *Business Strategy Review, 7*(4), 7-12. doi:10.1111/j.1467-8616.1996.tb00136.x

Barnett, W. P., & Hansen, M. T. (1996). The Red Queen in organizational evolution. *Strategic Management Journal, 17*, 139-157. doi:10.1002/smj.4250171010

Bastedo, M. N., & Bowman, N. A. (2010). *U.S. News & World Report* college rankings: Modeling institutional effects on organizational reputation. *American Journal of Education, 116*(2), 163-183. Retrieved from http://www.jstor.org/stable/10.1086/649437

Bastedo, M. N., & Bowman, N. A. (2011). College rankings as an interorganizational dependency: Establishing the foundation for strategic and institutional accounts. *Research in Higher Education, 52*(1), 3-23. doi:10.1007/s11162-010-9185-0

Beauducel, A., & Herzberg, P. Y. (2006). On the performance of maximum likelihood versus means and variance adjusted weighted least squares estimation in CFA. *Structural Equation Modeling: A Multidisciplinary Journal, 13*(2), 186–203. doi:10.1207/s15328007sem1302_2

Belsley, D. A., Kuh, E., & Welsh, R. E. (1980). *Regression diagnostics: Identifying influential data and sources of collinearity.* New York, NY: John Wiley & Sons, Inc.

Bennis, W. G., & O'Toole, J. (2005, May). How business schools lost their way. *Harvard Business Review, 38*(5), 96-104.

Bentler, P. M., & Chou, C. P. (1987). Practical issues in structural modeling. *Sociological Methods & Research, 16*(1), 78-117. doi:10.1177/0049124187016001004

Berger, M. (2001). Why the *U.S. News and World Report* law school rankings are both useful and important. *Journal of Legal Education, 51*(4), 487-502.

Bickerstaffe, G., & Ridgers, B. (2007). Ranking of business schools. *Journal of Management Development, 26*(1), 61-66. doi:10.1108/02621710710720103

Biesanz, J. C., & West, S. G. (2004). Towards understanding assessments of the Big Five: Multitrait-multimethod analyses of convergent and discriminant validity across measurement occasion and type of observer. *Journal of Personality, 72*(4), 845-876. doi:10.1111/j.0022-3506.2004.00282.x

Blackburn, G. (2011). Which Master of Business Administration (MBA)? Factors influencing prospective students' choice of MBA programme – an empirical study. *Journal of Higher Education Policy and Management, 33*(5), 473-483. doi:10.1080/1360080X.2011.605222

Blalark, F. J. (2012). *Utilizing principal-agent theory and data envelopment analysis to examine efficiency of resource utilization in undergraduate education for public and private non-profit four-year research universities* (Doctoral dissertation, the University of Minnesota). Retrieved from http://purl.umn.edu/136257

*Bloomberg Businessweek*. (1998, October 19). *Crunching the numbers: A tale of two surveys*. Retrieved from www.businessweek.com/datedtoc/1998/981019.htm

*Bloomberg Businessweek*. (2004, October 18). *A grads' – eye view*. pp. 80-81. Retrieved from http://www.businessweek.com/pdfs/2004/0442_bschoolstats.pfd

*Bloomberg Businessweek*. (2011). *The best U.S. B-schools of 2010.* Retrieved from http://www.businessweek.com/interactive_reports/bs_2010_US_FTMBA_TAB_1111.html

Boomsma, A., & Hoogland, J. J. (2001). The robustness of LISREL modeling revisited. In R. Cudeck, S. du Toit, & D. Sörbom (Eds.), *Structural equation modeling:*

*Present and future. A Festschrift in honor of Karl Jöreskog* (pp. 139-168). Lincolwood, IL: Scientific Software Intranational, Inc.

Bornstein, R. F. (2004). Face validity. In M. S. Lewis-Beck, A. Bryman, & T. F. Liao (Eds.), *Encyclopedia of social science research methods.* Thousand Oaks, CA: SAGE Publications, Inc. doi:http://dx.doi.org/10.4135/9781412950589.n323

Boruch, R. F., & Wolins, L. (1970). A procedure for estimation of trait, method, and error variance attributable to a measure. *Educational and Psychological Measurement, 30*(3), 547-574. doi:10.1177/001316447003000304

Boruch, R. F., Larkin, J. D., Wolins, L., & MacKinney, A. C. (1970). Alternative methods of analysis: Multitrait-multimethod data. *Educational and Psychological Measurement, 30*(4), 833-853. doi:10.1177/001316447003000405

Bowman, N. A., & Bastedo, M. N. (2009). Getting on the first page: Organizational reputation, status signals, and the impact of *U.S. News and World Report* on student decisions. *Research in Higher Education, 50*, 415-436. doi:10.1007/s11162-009-9129-8

Bowman, N. A., & Bastedo, M. N. (2011). Anchoring effects in world university rankings: Exploring biases in reputation scores. *Higher Education, 61*(4), 431-444. doi:10.1007/s10734-010-9339-1

Box, G. E. P., & Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American Statistical Association, 65*(332), 1509-1526. doi:10.2307/2284333

Bradshaw, D. (2007). Business school rankings: The love-hate relationship. *Journal of Management Development, 26*(1), 54-60. doi:10.1108/02621710710720095

Brooks, R. (2005). Measuring university quality. *The Review of Higher Education,* 29(1), 1-21. doi:10.1353/rhe.2005.0061

Brown, T. A. (2006). *Confirmatory factor analysis for applied research.* New York, London: The Guilford Press.

Brownstein, B. P. (1980). Pareto optimality, external benefits and public goods: A subjectivist approach. *The Journal of Libertarian Studies, 4*(1), 93-106.

Burke, J. C. (Ed.). (2005). *Achieving accountability in higher education: Balancing public, academic, and market demands*. San Francisco, CA: Jossey-Bass.

Byrne, B. M. (1994). *Structural equation modeling with EQS and EQS/Windows.* Thousand Oaks, CA: Sage Publications.

Byrne, B. M. (1998). *Structural equation modeling with LISREL, PRELIS and SIMPLIS: Basic concepts, applications and programming.* Mahwah, NJ: Lawrence Erlbaum Associates.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*(2), 81-105. doi:10.1037/h0046016

Campbell, D. T., & O'Connell, E. J. (1967). Method factors in multitrait-multimethod matrices: Multiplicative rather than additive? *Multivariate Behavioral Research, 2*(4), 409-426. doi:10.1207/s15327906mbr0204_1

Carlson School of Management, University of Minnesota. (2010). *China executive MBA (CHEMBA): Lingnan (University) College - Guangzhou, China.* Retrieved from http://www.csom.umn.edu/global-executive-mba/partner-schools/china/index.html

Carlson, O. (2014). *Internationalization of the MBA curriculum and its impact on building students' global competences* (Doctoral dissertation, University of Minnesota-Twin Cities). Retrieved from http://conservancy.umn.edu/bitstream/handle/11299/167714/Carlson_umn_0130E_15342.pdf?sequence=1&isAllowed=y

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research, 2*(6), 429-444. doi:10.1016/0377-2217(78)90138-8

Cheslock, J. J., & Gianneschi, M. (2008). Replacing state appropriations with alternative revenue sources: The case of voluntary support. *The Journal of Higher Education, 79*(2), 208-229. doi:10.1353/jhe.2008.0012

Choi, J., Peters, M., & Mueller, R. O. (2010). Correlational analysis of ordinal data: From Pearson's *r* to Bayesian polychoric correlation. *Asia Pacific Education Review, 11*(4), 459-466. doi:10.1007/s12564-010-9096-y

Clarke, M. (2002). Some guidelines for academic quality rankings. *Higher Education in Europe, 27*(4), 443-459. doi:10.1080/0379772022000071922

Colbert, A., Levary, R. R., & Shaner, M. C. (2000). Determining the relative efficiency of MBA programs using DEA. *European Journal of Operational Research, 125*(3), 656-669. doi:10.1016/S0377-2217(99)00275-1

Cole, D. A. (1987). Utility of confirmatory factor analysis in test validation research. *Journal of Consulting and Clinical Psychology, 55*(4), 584-594. doi:10.1037/0022-006X.55.4.584

Cole, J. R., & Lipton, J. A. (1977). The reputations of American medical schools. *Social Forces, 55*(3), 662-684.

Collet, F., & Vives, L. (2013). From preeminence to prominence: The fall of U.S. business schools and the rise of European and Asian business schools in the *Financial Times* Global MBA Rankings. *Academy of Management Learning and Education, 12*(4), 540-563. doi:10.5465/amle.2011.0094

Cook, W. D., & Seiford, L. M. (2009). Data envelopment analysis (DEA) - Thirty years on. *European Journal of Operational Research, 192*, 1-17. doi:10.1016/j.ejor.2008.01.032

Corley, K., & Gioia, D. (2000). The rankings game: Managing business school reputation. *Corporate Reputation Review, 3*(4), 319-333.doi:10.1057/palgrave.crr.1540123

Cornuel, E. (2005). The role of business schools in society. *Journal of Management Development, 24*(9), 819-829. doi:10.1108/02621710510621321

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology, 78*(1), 98-104. doi:10.1037/0021-9010.78.1.98

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*(4), 281-302.

Dahlin-Brown, N. (2005). The perceptual impact of *U.S. News & World Report* rankings on eight public MBA programs. *Journal of Marketing for Higher Education, 15*(2), 155-179. doi:10.1300/J050v15n02_07

De Onzoño, S. I., & Carmona, S. (2007). The changing business model of B-schools. *Journal of Management Development, 26*(1), 22-32. doi:10.1108/02621710710720059

DeAngelo, H., DeAngelo, L., & Zimmerman, J. L. (2005, July). *What's really wrong with U.S. business schools?* Unpublished manuscript. University of Southern California. Retrieved from http://fisher.osu.edu/~young_53/DeAngelo-DeAngelo-Zimmerman.pdf

Denning, S. (2014, February 07). Why aren't business schools more business-like? *Forbes*. Retrieved from http://www.forbes.com/sites/stevedenning/2014/02/07/why-arent-business-schools-more-business-like/

Devinney, T., Dowling, G. R., & Perm-Ajchariyawong, N. (2008). The *Financial Times* business schools ranking: What quality is this signal of quality? *European Management Review, 5*(4), 195-208. doi:10.1057/emr.2008.14

Dichev, I. D. (1999). How good are business school rankings? *Journal of Business, 72*(2), 201-213.

Dichev, I. D. (2008). Comment: The *Financial Times* business schools ranking: What quality is this signal of quality? *European Management Review, 5*(4)*,* 219-224. doi:10.1057/emr.2008.23

Dodge, Y. (2010). *The concise encyclopedia of statistics*. New York, NY: Springer.

Dowling, J., & Pfeffer, J. (1975). Organizational legitimacy: Social values and organizational behavior. *The Pacific Sociological Review, 18*(1), 122-136.

Drnevich, P. L., Armstrong, C. E., Crook, T. A., & Crook, T. R. (2011). Do research and education matter to business school rankings? *International Journal of Management in Education, 5*(2/3), 169-187. doi:10.1504/IJMIE.2011.039483

Eberhardt, B. J., & Moser, S. (1997). Business concerns regarding MBA education: Effects on recruiting. *Journal of Education for Business, 72*(5), 293-296.

Eckles, J. E. (2010). Evaluating the efficiency of top liberal arts colleges. *Research in Higher Education, 51*, 266–293. doi:10.1007/s11162-009-9157-4

Eisenhardt, K. M., & Bourgeois, L. J., III. (1988). Politics of strategic decision making in high-velocity environments. *Academy of Management Review, 31*(4), 737-771.

Ekström, J. (2011). *On the relation between the polychoric correlation coefficient and Spearman's rank correlation coefficient.* (Department of Statistics Papers). UCLA: Department of Statistics. Retrieved from http://escholarship.org/us/item/7j01t5sf

Elliott, C., & Soo, K. T. (2013). The international market for MBA qualifications: The relationship between tuition fees and applications. *Economics of Education Review, 34*, 162-174. doi:10.1016/j.econedurev.2013.02.006

Elsbach, K. D., & Kramer, R. M. (1996). Members' responses to organizational identity threats: Encountering and countering the business week rankings. *Administrative Science Quarterly, 41*(3), 442-476.

Espeland, W. N., & Sauder, M. (2007). Rankings and reactivity: How public measures recreate social worlds. *American Journal of Sociology, 113*(1), 1-40.

EUA Annual Conference (2013, April 11-12). European Universities – Global Engagement. Ghent University, Ghent, Belgium.

Everitt, L. (2014). MBA rankings and B-schools: A love-hate relationship. *Poets & Quants, 1*(2), 14-17. Retrieved from http://poetsandquants.com/wp-content/uploads/2014/05/PQ2_FULLMAG.pdf

Ewell, P. T. (1994). A matter of integrity: Accountability and the future of self-regulation. *Change, 26*(6), 24-29.

Fee, C. E., Hadlock, C. J., & Pierce, J. R. (2005). Business school rankings and business school deans: A study of nonprofit governance. *Financial Management, 34*(1), 143-166.doi:10.1111/j.1755-053X.2005.tb00095.x

Fisher, D. M., Kiang, M., & Fisher, S. A. (2007). A value-added approach to selecting the best master of business administration (MBA) program. *Journal of Education for Business, 83*(2), 72-76.

Flanigan, S., & Morse, R. (2015, March 9). *Methodology: 2016 Best Business Schools rankings.* Retrieved from http://www.usnews.com/education/best-graduate-schools/articles/business-schools-methodology

Flora, D. B., & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for Confirmatory Factor Analysis with ordinal data. *Psychological Methods, 9*(4), 466-491. doi:10.1037/1082-989X.9.4.466

Fombrun, C. J. (1996). *Reputation: Realizing value from corporate image*. Boston, MA: Harvard Business School Press.

Fombrun, C., & Shanley, M. (1990). What's in a name? Reputation-building and corporate strategy. *Academy of Management Journal, 33*(2), 233-258.

*Forbes.* (2011). *The best business schools.* Retrieved from http://www.forbes.com/lists/2011/95/best-business-schools-11_rank.htm

Fox, J. (2016). *Applied regression analysis and generalized linear models* (3rd ed.). Thousand Oaks, CA: SAGE Publications, Inc.

Friga, P. N., Bettis, R. A., & Sullivan, R. S. (2003). Changes in graduate management education and new business school strategies for the 21st century. *Academy of Management Learning and Education, 2*(3), 233-249.

Furnham, A., & Boo, H. C. (2011). A literature review of the anchoring effect. *The Journal of Socio-Economics, 40*, 35-42. doi:10.1016/j.socec.2010.10.008

Gaddis, P. O. (2000). Business schools: Fighting the enemy within. *Strategy and Business, 21*(4). Retrieved from http://www.strategy-business.com/article/14038?gko=abff4

Gioia, D. A., & Corley, K. G. (2002). Being good versus looking good: Business school rankings and the Circean transformation from substance to image. *Academy of Management Learning and Education, 1*(1), 107-120. doi:10.5465/AMLE.2002.7373729

Gladwell, M. (2011, February 14). The order of things. *The New Yorker,* pp. 68-75.

Glick, W. (2008). Rain man or pied piper? Moving business schools beyond media rankings with mass customization and stakeholder education. *Academy of Management Perspectives, 22*(1), 18-23. doi:10.5465/AMP.2008.31217509

Gloeckler, G. (2006, June 6). Stanford's new-look MBA. *Business Week.* Retrieved from http://www.businessweek.com/bschools/content/jun2006/bs2006065_4760_bs001.htm

Gnolek, S. L., Falciano, V. T., & Kuncl, R. W. (2014). Modeling change and variation in *U.S. News & World Report* College Rankings: What would it really take to be in the top 20? *Research in Higher Education, 55*(8), 761-779. doi:10.1007/s11162-014-9336-9

Golding S. L., & Seidman, E. (1974). Analysis of multitrait-multimethod matrices: A two step principal components procedure. *Multivariate Behavioral Research, 9*(4), 479-496. doi:10.1207/s15327906mbr0904_7

Goldstein, M. J., Lunn, M. R., & Peng, L. (2015). What makes a top research medical school? A call for a new model to evaluate academic physicians and medical school performance. *Academic Medicine, 90*(5), 603-608. doi:10.1097/ACM.0000000000000646

Gourman, J. (1997). *The Gourman Report: Graduate programs*. New York, NY: The Princeton Review Publishing.

Graduate Management Admission Council (2012a). *Corporate recruiters survey.* Retrieved from http://www.gmac.com/~/media/Files/gmac/Research/Employment%20Outlook/2012-crs-report-final-participants.pdf

Graduate Management Admission Council (2012b). *The mba.com prospective students survey.* Retrieved from http://www.gmac.com/~/media/Files/gmac/Research/prospective-student-data/2012prospectivestudentssr.pdf

Greene, W. H. (2012). *Econometric analysis: International edition* (7th ed.). Boston, MA: Pearson Education Limited.

Guilford, J. P. (1954). *Psychometric methods* (2nd ed.). New York, NY: McGraw-Hill, Inc.

Gujarati, D. N. (1995). *Basic econometrics* (3rd ed.). New York, NY: McGraw-Hill, Inc.

Gwet, K. L. (2011). *On the Krippendorff's alpha coefficient.* Retrieved from http://www.agreestat.com/research_papers/onkrippendorffalpha.pdf

Haksever, C., & Muragishi, Y. (1998). Measuring the value in MBA programmes. *Education Economics, 6*(1), 11-25. Retrieved from http://web.b.ebscohost.com/ehost/detail/detail?sid=e55c138e-9dff-4240-bbec-b13612bf23c2%40sessionmgr120&vid=0&hid=115&bdata=JkF1dGhUeXB1PW1wLHVpZCZzaXR1PWVob3N0LWxpdmU%3d#AN=779664&db=keh

Harvard Business School. (2010). *Executive education: Custom programs.* Retrieved http://www.exed.hbs.edu/programs/custom/Pages/default.aspx

Harvard will launch MD/MBA program. (2004, May 24). *Modern Healthcare, 34*(21). Retrieved from http://web.ebscohost.com.floyd.lib.umn.edu/ehost/detail?hid=12&sid=51ff5131-edd2-4f22-aed3-7f8781ea9429%40sessionmgr11&vid=10&bdata=JnNpdGU9ZWhvc3QtbGl2ZQ%3d%3d#db=aph&AN=13404410

Harwell, M. R., & Gatti, G. G. (2001). Rescaling ordinal data to interval data in educational research. *Review of Educational Research, 1*(71), 105-131. Retrieved from http://www.jstor.org/stable/3516068

Hawawini, G. (2005). The future of business schools. *Journal of Management Development, 24*(9), 770-782. doi:10.1108/026217105106221286

Haworth, D. (2013, April 20). Pressing need for more sophisticated rankings – EUA. *University World News, 268*. Retrieved from http://www.universityworldnews.com/article.php?story=20130418131705959

Hay, M. (2008). Business schools: A new sense of purpose. *Journal of Management Development,* 27(4), 371-378. doi:10.1108/02621710810866723

Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, *1*(1), 77-89. doi:10.1080/19312450709336664

Hayward, M., & Boeker, W. (1998). Power and conflicts of interest in professional firms: Evidence from investment banking. *Administrative Science Quarterly, 43*(1), 1-22.

Hensel, R., Meijers, F., van der Leeden, R., & Kessels, J. (2010). 360 degree feedback: How many raters are needed for reliable ratings on the capacity to develop competences, with personal qualities as developmental goals? *The International Journal of Human Resource Management, 21*(15), 2813-2830. doi:10.1080/09585192.2010.528664

Holbrook, M. B. (2007). Objective characteristics, subjective evaluations, and possible distorting biases in the business-school rankings: The case of *U.S. News & World Report*. *Marketing Education Review, 17*(2), 1-12.

Hopkins, K. (2012, August 28). 10 MBAs that lead to low student debt: For students who borrow, these business schools could be a bargain. *U.S. News & World Report.* Retrieved from http://www.usnews.com/education/best-graduate-schools/the-short-list-grad-school/articles/2012/08/28/10-mbas-that-lead-to-low-student-debt

Hou, A. Y.-C., Morse, R., & Shao, Y. E. (2012). Is there a gap between students' preference and university presidents' concern over college ranking indicators?: A case study of "College Navigator in Taiwan". *Higher Education, 64*, 767-787. doi:10.1007/s10734-012-9524-5

Howell, D. C. (2007). *Statistical methods for psychology* (6th ed.). Belmont, CA: Thomson Wadsworth.

Hox, J. J., Maas, C. J. M., & Brinkhuis, M. J. S. (2010). The effect of estimation method and sample size in multilevel structural equation modeling. *Statistica Neerlandica, 64*(2), 157-170. doi:10.1111/j.1467-9574.2009.00445.x

Hoyle, R. H. (Ed.). (1995). *Structural equation modeling: Concepts, issues, and applications.* Thousand Oaks, CA: SAGE Publications, Inc.

Hsu, M. K., Chao, G. H., & James, M. L. (2009). An efficiency comparison of MBA programs: Top 10 versus non-top 10. *Journal of Education for Business, 84*(5), 269-274. doi:10.3200/JOEB.84.5.269-274

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1-55. doi:10.1080/10705519909540118

Hubbard, G. (2015, April 6). Do B-school rankings really matter? *Fortune.* Retrieved from http://fortune.com/2015/04/06/do-b-school-rankings-really-matter/

Hubert, L. J., & Baker, F. B. (1978). Analyzing the multitrait-multimethod matrix. *Multivariate Behavioral Research, 13*(2), 163-179. doi:10.1207/s15327906mbr1302_3

Iacobucci, D. (2010). Structural equations modeling: Fit indices, sample size, and advanced topics. *Journal of Consumer Psychology, 20*(1), 90–98. doi:10.1016/j.jcps.2009.09.003

Iacobucci, D. (2013). A psychometric assessment of the *Businessweek, U.S. News & World Report,* and *Financial Times* rankings of business schools' MBA programs. *Journal of Marketing Education, 35*(3), 204-219. doi:10.1177/0273475313491576

INSEAD. (2010). *MBA programme. North & South America events.* Retrieved from http://mba.insead.edu/events/north_south_america.cfm

Institute for Higher Education Policy. (2009, September). *The role and relevance of rankings in higher education policymaking* (Issue Brief). Washington, DC: Sponsler. Retrieved from http://www.ihep.org/sites/default/files/uploads/docs/pubs/issue_brief_the_role_and_relevance_of_rankings.pdf

Jaccard, J., & Wan, C. K. (1996). *LISREL approaches to interaction effects in multiple regression.* Thousand Oaks, CA: Sage Publications.

Jackson, D. N. (1969). Multimethod factor analysis in the evaluation of convergent and discriminant validity. *Psychological Bulletin, 72*(1), 30-49. doi:10.1037/h0027421

John, E. P. St., Kline, K. A., & Asker, E. H. (2001). The call for public accountability: Rethinking the linkages to student outcomes. In D. E. Heller (Ed.), *The states and public higher education policy: Affordability, access, and accountability* (pp. 219-242). Baltimore and London: The Johns Hopkins University Press.

Jöreskog, K. G., & Sörbom, D. (2001). *LISREL 8 user's reference guide*. Chicago, IL: Scientific Software International, Inc.

Jöreskog, K. G., & Sörbom, D. (2014). *LISREL 9.10* [Computer software]. Chicago, IL: Scientific Software International, Inc. Retrieved from http://www.ssicentral.com

Jöreskog, K. G., Sörbom, D., Du Toit, S. H. C. & Du Toit, M. (2001). *LISREL 8: New statistical features* (3rd ed.). Lincolnwood, IL: Scientific Software International, Inc.

Kalleberg, A. L., & Kluegel, J. R. (1975). Analysis of the multitrait-multimethod matrix: Some limitations and an alternative. *Journal of Applied Psychology, 60*(1), 1-9. doi:10.1037/h0076267

Kenny, D. A., & Kashy, D. A. (1992). Analysis of the multitrait-multimethod matrix by confirmatory factor analysis. *Psychological Bulletin, 112*(1), 165-172. doi:10.1037/0033-2909.112.1.165

Kleiber, C., & Zeileis, A. (2008). *Applied econometrics with R.* New York, NY: Springer Science+Business Media, LLC.

Klein, S. P., & Hamilton, L. (1998). *The validity of the U.S. News and World Report ranking of ABA law schools.* Retrieved from The Association of American Law Schools website: http://www.aals.org/reports/validity.html

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York, NY: The Guilford Press.

Korobkin, R. (1998). In praise of law school rankings: Solutions to coordination and collective action problems. *Texas Law Review, 77*, 403-428.

Krause, M. S. (1972). The implications of convergent and discriminant validity data for instrument validation. *Psychometrika, 37*(2), 179-186. doi:10.1007/BF02306775

Krippendorff, K. (2004a). *Content analysis: An introduction to its methodology* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Krippendorff, K. (2004b). Reliability in content analysis: Some common misconceptions and recommendations. *Human Communication Research, 30*(3), 411-433. doi:10.1111/j.1468-2958.2004.tb00738.x

Krippendorff, K. (2011). *Computing Krippendorff's alpha-reliability.* Retrieved from http://repository.upenn.edu/cgi/viewcontent.cg?article=1043Lcontext=asc.papers

Kutner, M. H., Nachtshein, C. J., Neter, J., & Li, W. (2005). *Applied linear statistical models* (5th ed.). New York, NY: McGraw-Hill/Irwin.

Langer, D. A., Wood, J. J., Bergman, R. L., & Piacentini, J. C. (2010). A multitrait-multimethod analysis of the construct validity of child anxiety disorders in a clinical sample. *Child Psychiatry and Human Development, 41*(5), 549-561. doi:10.1007/s10578-010-0187-0

Lavelle, L. (2012a, November 15). Best b-schools ranking history. *Bloomberg Businessweek.* Retrieved from www.businessweek.com/articles/2012-11-15/best-b-schools-ranking-history

Lavelle, L. (2012b, November 15). How we ranked the b-schools. *Bloomberg Businessweek*. Retrieved from www.businessweek.com/articles/2012-11-15/how-we-ranked-the-b-schools

Leibert, A. (2014, April). Introducing the Graduate Employment Outcomes tool. *Minnesota Employment Review.* Retrieved from http://mn.gov/deed/newscenter/publications/review/April-2014/college-education-pay.jsp

Levene, H. (1960). Robust tests for equality of variances. In I. Olkin (Ed.), *Contributions to probability and statistics* (pp. 278-292). Palo Alto, CA: Stanford University Press.

Lewin, A. Y., Morey, R. C., & Cook, T. J. (1982). Evaluating the administrative efficiency of courts. *Omega: The International Journal of Management Science, 10*(4), 401-411. doi:10.1016/0305-0483(82)90019-6

Liberto, J. (2013, August 22). Obama proposes new college ratings system. *CNNMoney.* Retrieved from http://money.cnn.com/2013/08/22/pf/college/obama-college-ratings/index.html?iid=EL

Liebowitz, S. J. (2003). *The role of research in business school rankings* (Working Paper). University of Texas at Dallas: Department of Finance & Managerial Economics. Retrieved from http://ssrn.com/abstract=464480

Ljung, G. M., & Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika, 65*(2), 297-303. doi:10.2307/2335207

Lorange, P. (2005). Strategy means choice: Also for today's business school! *Journal of Management Development, 24*(9), 783-790. doi:10.1108/02621710510621295

Luo, H, (2011). *Some aspects of confirmatory factor analysis of ordinal variables and generating non-normal data* (Doctoral dissertation, Uppsala University, Uppsala, Sweden). Retrieved from Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences http://uu.diva-portal.org/smash/record.jsf?pid=diva2:405108

Lytle, R. (2012, April 24). 10 Business schools that lead to most debt: Some programs leave students with nearly $100,000 in debt at graduation. *U.S. News & World Report.* Retrieved from http://www.usnews.com/education/best-graduate-schools/the-short-list-grad-school/articles/2012/04/24/10-business-schools-that-lead-to-most-debt

MacCallum, R. C., Widaman, K. F., Preacher, K. J., & Hong, S. (2001). Sample size in factor analysis: The role of model error. *Multivariate Behavioral Research, 36*(4), 611-637. doi:10.1207/S15327906MBR3604_06

Marginson, S., & van der Wende, M. (2007). To rank or to be ranked: The impact of global rankings in higher education. *Journal of Studies in International Education, 11*(3/4), 306-329. doi:10.1177/1028315307303544

Marsh, H. W. (1989). Confirmatory factor analyses of multitrait-multimethod data: Many problems and a few solutions. *Applied Psychological Measurement, 13*(4), 335-361. doi:10.1177/014662168901300402

Marsh, H. W., & Grayson, D. (1995). Latent variable models of multitrait-multimethod data. In R. H. Hoyle (Ed.). *Structural equation modeling: Concepts, issues, and applications* (pp. 177-198). Thousand Oaks, CA: SAGE Publications, Inc.

Martins, L. L. (2005). A model of the effects of reputational rankings on organizational change. *Organization Science, 16*(6), 701-720. doi:10.1287/orsc.1050.0144

Maruyama, G. M. (1998). *Basics of structural equation modeling.* Thousand Oaks, CA: SAGE Publications, Inc.

McDonough, P. M., Antonio, A. L., Walpole, M. B., & Perez, L. X. (1998). College rankings: Democratized college knowledge for whom? *Research in Higher Education, 39*(5), 513-537. doi:10.1023/A:1018797521946

McGuire, M. D. (1995). Validity issues for reputational studies. *New Directions for Institutional Research, 1995*(88), 45-59. doi:10.1002/ir.37019958806

Meredith, M. (2004). Why do universities compete in the ratings game? An empirical analysis of the effects of the *U.S. News and World Report* college rankings. *Research in Higher Education, 45*(5), 443-461. doi:10.1023/B:RIHE.0000032324.46716.f4

Merisotis, J., & Sadlak, J. (2005). Higher education rankings: Evolution, acceptance, and dialogue. *Higher Education in Europe, 30*(2), 97-101. doi:10.1080/03797720500260124

Merton, R. K. (1968). The Matthew effect in science. *Science, 159*(3810), 56-63.

Mintzberg, H. (1996, July-August). Musing on management: Ten ideas designed to rile everyone who cares about management. *Harvard Business Review, 74*(4), 61-67.

Mintzberg, H. (2004). *Managers, not MBAs: A hard look at the soft practice of managing and management development*. San Francisco, CA: Berrett-Koehler Publishers.

Mirabella, R. M. (2007). University-based educational programs in nonprofit management and philanthropic studies: A 10-year review and projections of future trends. *Nonprofit and Voluntary Sector Quarterly*, Supplement *36*(4), 11S-27S. doi:10.1177/0899764007305051

Monks, J., & Ehrenberg, R. G. (1999). *The impact of U.S. News and World Report college rankings on admission outcomes and pricing policies at selective private institutions.* (NBER Working Paper No. 7227). Cambridge, MA: National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w7227.pdf

Montgomery, D. C., Jennings, C. L., & Kulanchi, M. (2008). *Introduction to time series analysis and forecasting*. Hoboken, NJ: John Wiley & Sons Publication.

Morgeson, F. P., & Nahrgang, J. D. (2008). Same as it ever was: Recognizing stability in the *Business Week* rankings. *Academy of Management Learning and Education, 7*(1), 26-41. doi:10.5465/AMLE.2008.31413860

Morse, R., & Flanigan, S. (2011). Business school rankings methodology. *U.S. News & World Report.* Retrieved from http://www.usnews.com/education/best-graduate-schools/articles/2011/03/14/business-school-rankings-methodology-2012

Nardi, P. M. (2006). *Doing survey research: A guide to quantitative methods* (2nd ed). Boston, MA: Pearson Education, Inc.

National Science Foundation/National Center for Science and Engineering Statistics. (2013). *Higher education research and development: Fiscal Year 2011.* Retrieved from www.nsf.gov/statistics/nsf13325/pdf/nsf13325.pdf

O'Connell, C. (2013). Research discourses surrounding global university rankings: Exploring the relationship with policy and practice recommendations. *Higher Education: The International Journal of Higher Education and Educational Planning, 65*(6), 709-723. doi:10.1007/s10734-012-9572-x

Olsson, U. H., Foss, T., Troye, S. V., & Howell, R. D. (2000). The performance of ML, GLS, and WLS estimation in structural equation modeling under conditions of misspecification and nonnormality. *Structural Equation Modeling: A Multidisciplinary Journal, 7*(4), 557-595. doi:10.1207/S15328007SEM0704_3

Osborn, J. W. (2010). Improving your data transformations: Applying the Box-Cox transformation. *Practical Assessment, Research & Evaluation, 15*(12). Retrieved from http://pareonline.net/getvn.asp?v=15&n=12

Peters, K. (2007). Business school rankings: Content and context. *Journal of Management Development, 26*(1), 49-53. doi:10.1108/0262171071070086

Pfeffer, J., & Fong, C. T. (2002). The end of business schools? Less success than meets the eye. *Academy of Management Learning and Education, 1*(1), 78-95. doi:10.5465/AMLE.2002.7373679

Pfeffer, J., & Fong, C. T. (2003). Assessing business schools: Reply to Connolly. *Academy of Management Learning and Education, 2*(4), 368-370. Retrieved fromhttp://www.jstor.org/stable/40214214

Pfeffer, J., & Fong, C. T. (2004). The business school 'business': Some lessons from the US experience. *Journal of Management Studies, 41*(8), 1501-1520.

Pfeffer, J., & Salancik, G. R. (1978). *The external control of organizations: A resource dependence perspective*. New York, NY: Harper & Row.

Pitsakis, K., Souitaris, V., & Nicolaou, N. (2015). The peripheral halo effect: Do academic spinoffs influence universities' research income? *Journal of Management Studies, 52*(3), 321-353. doi:10.1111/joms.12119

Podolny, J. M. (1993). A status-based model of market competition. *The American Journal of Sociology, 98*(4), 829-872.

Podolny, J. M. (2009, June). The buck stops (and starts) at business school. *Harvard Business Review, 87*(6), 62-67.

Policano, A. J. (2001). Ten easy steps to a top-25 MBA program. *Selections, 1*(2), 39-40.

Policano, A. J. (2007). The rankings game: And the winner is … *Journal of Management Development, 26*(1), 43-48. doi:10.1108/02621710710720077

Popping, R. (1988). On agreement indices for nominal data. In W. E. Saris & I. N. Gallhofer (Eds.), *Sociometric research: Data collection and scaling* (pp. 90–105). Cambridge, MA: MIT Press.

Proulx, R. (2009). World university rankings. The need for a new paradigm. In B. M. Kehm & B. Stensaker (Eds.), *University rankings, diversity, and the new landscape of higher education* (pp. 35-46). Rotterdam, The Netherlands: Sense Publishers.

R version 3.1.0 (2014-04-10) [Computer software]. The R Foundation for Statistical Computing.

Rao, H. (1994). The social construction of reputation: Certification contests, legitimation, and the survival of organizations in the American automobile industry: 1895-1912. *Strategic Management Journal, 15*, 29-44.

Rauhvargers, A. (2013, April 11-12). *Impact of rankings is growing: What should universities do?* Paper presented at the EUA Annual Conference 2013, Ghent University, Ghent, Belgium. Retrieved from

http://www.eua.be/events/past/2013/EUA-Annual-Conference-2013/Presentations_and_photos.aspx

Reichardt, C. S., & Coleman, S. C. (1995). The criteria for convergent and discriminant validity in a multitrait-multimethod matrix. *Multivariate Behavioral Research, 30*(4), 513-538. doi:10.1207/s15327906mbr3004_3

Rindova, V., Williamson, I., Petkova, A., & Sever, J. M. (2005). Being good or being known: An empirical examination of the dimensions, antecedents, and consequences of organizational reputation. *The Academy of Management Journal, 48*(6), 1033-1049.

Roberts, P. W., & Dowling, G. R. (2002). Corporate reputation and sustained superior financial performance. *Strategic Management Journal, 23*(12), 1077-1093.

Röbken, H. (2004). *Inside the „knowledge factory": Organizational change in business schools in Germany, Sweden and the USA.* Weisbaden, Germany: Deutscher Universitäts-Verlag. doi:10.1007/978-3-322-81180_6_1

Rodriguez, M. C., & Maeda, Y. (2006). Meta-analysis of coefficient alpha. *Psychological Methods, 11*(3), 306-322. doi:10.1037/1082-989X.11.3.306

Rudolph, F. (1962). *The American college and university: A history.* New York, NY: Knopf.

Safón, V. (2007). Factors that influence recruiters' choice of B-schools and their MBA graduates: Evidence and implications for B-schools. *Academy of Management Learning and Education, 6*(2), 217-233.

Safón, V. (2009). Measuring the reputation of top US business schools: A MIMIC modeling approach. *Corporate Reputation Review, 12*(3), 204-228. doi:10.1057/crr.2009.19

Safón, V. (2012). Can the reputation of an established business school change? *Management in Education, 26*(4), 169-180. doi:10.1177/0892020611433081

Samy, M., Hodgson, C., & Moulton, D. (2007). An examination of the *Economist* Intelligence Unit's ranking and its application to curriculum design of an MBA programme. *International Journal of Management Education, 6*(2), 79-94.

Sauder, M., & Espeland, W. N. (2006). Strength in numbers? The advantages of multiple rankings. *Indiana Law Journal, 81*(1), 205-227.

Sauder, M., & Lancaster, R. (2006). Do rankings matter? The effects of the *U.S. News & World Report* rankings on the admissions process of Law schools. *Law & Society Review, 40*(1), 105-134.

Schatz, M., & Crummer, R. E. (1993). What's wrong with MBA ranking surveys? *Management Research News, 16*(7), 15-18. Retrieved from http://officialmbaguide.org/whatswrong.php

Schmalbeck, R. (1998). The durability of law school reputation. *Journal of Legal Education, 48*(4), 568-590.

Schmitt, N., & Stults, D. M. (1986). Methodology review: Analysis of multitrait-multimethod matrices. *Applied Psychological Measurement, 10*(1), 1-22. doi:10.1177/014662168601000101

Schriesheim, C. A. (1981a). The effect of grouping or randomizing items on leniency response bias. *Educational and Psychological Measurement, 41*(2), 401-411. doi:10.1177/001316448104100219

Schriesheim, C. A. (1981b). Leniency effects on convergent and discriminant validity for grouped questionnaire items: A further investigation. *Educational and Psychological Measurement, 41*(4), 1093-1099. doi:10.1177/001316448104100419

Schumacker, R. R., & Lomax, R. G. (2004). *A beginner's guide to structural equation modeling* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Seymour, D. T. (1992). *On Q: Causing quality in higher education*. New York, NY: American Council on Education and Macmillan.

Sheather, S. J. (2009). *A modern approach to regression with R.* New York, NY: Springer Science+Business Media, LLC.

Siemens, J. C., Burton, S., Jensen, T., & Mendoza, N. A. (2005). An examination of the relationship between research productivity in prestigious business journals and popular press business school rankings. *Journal of Business Research, 58*(4), 467-476. doi:10.1016/j.jbusres.2003.07.003

Silver, N. C., & Dunlap, W. P. (1987). Averaging correlation coefficients: Should Fisher's *z* transformation be used? *Journal of Applied Psychology, 72*(1), 146-148. doi:10.1037/0021-9010.72.1.146

Smith, K. T., Smith, M., & Wang, K. (2010). Does brand management of corporate reputation translate into higher market value? *Journal of Strategic Marketing, 18*(3), 201-221. doi:10.1080/09652540903537030

Social Science Research Network. (2016). *SSRN top 1,000 business schools.* Retrieved from http://www.ssrn.com/en/

Stake, J. E. (2006). The interplay between law school rankings, reputations, and resource allocation: Ways rankings mislead. *Indiana Law Journal, 81*(1), 229-270.

Stanley, J. C. (1961). Analysis of unreplicated three-way classifications, with applications to rater bias and trait independence. *Psychometrika, 26*(2), 205-2019. doi:10.1007/BF02289715

Stanley, T. (2010, August 09). The 10 most innovative business school classes. *Forbes.* Retrieved from http://www.forbes.com/2010/08/09/most-innovative-business-school-classes-entrepreneurs-management-sustainable-tech-10-innovative.html

Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research, 25*(2), 173-180. doi:10.1207/s15327906mbr2502_4

Steiger, J. H. (2007). Understanding the limitations of global fit assessment in structural equation modeling. *Personality and Individual Differences, 42*(5), 893-898. doi:10.1016/j.paid.2006.09.017

Stemler, S. E. (2004). A comparison of consensus, consistency, and measurement approaches to estimating interrater reliability. *Practical Assessment, Research & Evaluation, 9*(4). Retrieved from http://pareonline.net/getvn.asp?v=9&n=4

Stevens, J. P. (1984). Outliers and influential data points in regression analysis. *Psychological Bulletin, 95*(2), 334-344.

Stolz, I., Hendel, D. D., & Horn, A. S. (2010). Ranking of rankings: Benchmarking twenty-five higher education ranking systems in Europe. *Higher Education, 60*(5), 507-528. doi:10.1007/s10734-010-9312-z

Stratford, M. (2015, August 12). The non-ratings future. *Inside Higher Ed.* Retrieved from https://www.insidehighered.com/news/2015/08/12/ratings-fight-over-education-department-prepares-launch-new-consumer-tool

Strauss, M. E., & Smith, G. T. (2009). Construct validity: Advances in theory and methodology. *The Annual Review of Clinical Psychology, 5*, 1-25. doi:10.1146/annurev.clinpsy.032408.153639

Sweitzer, K., & Volkwein, J. F. (2009). Prestige among graduate and professional schools: Comparing the U.S. News' graduate school reputation ratings between disciplines. *Research in Higher Education, 50*(8), 812-836. doi:10.1007/s11162-009-9140-0

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Pearson/Allyn & Bacon.

Tanaka, J. S. (1987). "How big is big enough?": Sample size and goodness of fit in structural equation models with latent variables. *Child Development, 58*(1), 134-146. doi:10.2307/1130296

Tauer, L. W., Fried, H. O., & Fry, W. E. (2007). Measuring efficiencies of academic departments within a college. *Education Economics, 15*(4), 473–489. doi:10.1080/09645290701419516

The Accreditation Council for Business Schools and Programs (ACBSP). (2013). *About ACBSP*. *Discover ACBSP*. Retrieved from http://www.acbsp.org/?page=about_us

*The Economist* (2012a, March 13). *2011 Full Time MBA Ranking: United States.* Retrieved from http://www.economist.com/whichmba/full-time-mba-ranking?term_node_tid_depth=77633

*The Economist.* (2012b). *Which MBA? Methodology.* Retrieved from http://www.economist.com/whichmba/methodology

*The Financial Times* (2011). *Global MBA Rankings 2011.* Retrieved from http://rankings.ft.com/exportranking/global-mba-rankings-2011/pdf

*The Financial Times* (2015). *Methodology: FT Global MBA Ranking 2015.* Retrieved from http://www.ft.com/intl/cms/s/2/75fc5c24-9b1d-11e4-b651-00144feabdc0.html#axzz3svg5GUjm

The International Ranking Expert Group (IREG). (2006). *Berlin Principles on Ranking of Higher Education Institutions.* Berlin, Germany: The International Ranking Expert Group. Retrieved from http://www.che.de/doenloads/Berlin_Principles_IREG-534.pdf

The International Ranking Expert Group (IREG). (2011). *The IREG ranking audit manual.* Retrieved from http://www.ireg-observatory.org/pdf/ranking_audith_audit.pdf

The International Ranking Expert Group (IREG). (2015a). *Ranking audit.* Retrieved from http://ireg-observatory.org/en/index.php/information

The International Ranking Expert Group (IREG). (2015b). *IREG Inventory of national rankings.* Retrieved from http://ireg-observatory.org/en/index.php/ireg-inventory

TheLadders (2012, March 21). *Keeping an eye on recruiter behavior.* Retrieved from http://cdn.theladders.net/static/images/basicSite/pdfs/TheLadders-EyeTracking-StudyC2.pdf

Thomas, H., Lorange, P., & Sheth, J. (2013). *The business school in the twenty-first century: Emergent challengers and new business models.* New York, NY: Cambridge University Press.

Thomson Reuters (2010). *Global opinion survey: New outlooks on institutional profiles.* Retrieved from http://ip-science.thomsonreuters.com/m/pdfs/Global_Opinion_Survey.pdf

Thorndike R. M., & Thorndike-Christ, T. (2010). *Measurement and evaluation in psychology and education* (8th ed.). Boston, MA: Pearson Education, Inc.

Torraco, R. J. (2005). Organization development: A question of fit for universities. *Advances in Developing Human Resources, 7*(3), 303-310. doi:10.1177/1523422305277171

Torraco, R. J., & Hoover, R. E. (2005). Organization development and change in universities: Implications for research and practice. *Advances in Developing Human Resources, 7*(3), 422-437. doi:10.1177/1523422305277180

Tracy, J., & Waldfogel, J. (1997). The best business schools: A market-based approach. *Journal of Business, 70*(1), 1-31.

Tucker, L. R. (1966). Some mathematical notes on three-mode factor analysis. *Psychometrika, 31*(3), 279-311. doi:10.1007/BF02289464

Tucker, L. R. (1967). Three-mode factor analysis of Parker-Fleishman complex tracking behavior data. *Multivariate Behavioral Research, 2*(2), 139-151. doi:10.1207/s15327906mbr0202_1

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124-1131. Retrieved from http://www.jstor.org/stable/1738360

U.S. Department of Education, National Center for Education Statistics. (2011, April). *Digest of education statistics, 2010* (NCES 2011-015). Retrieved from http://nces.ed.gov/programs/digest/d10/ch_3.asp

*U.S. News & World Report.* (2011). *U.S. News & World Report best graduate schools guidebook* (2012 edition). Washington, D.C.: *U.S. News & World Report* L.P

*U.S. News & World Report*. (2013a). *Best Business Schools.* Retrieved from http://premium.usnews.com/best-graduate-schools/top-business-schools/mba-rankings/int%2Bgrad-best-business/sort%2Bc_avg_comp_base_bonus/sortdir%2Basc

*U.S. News & World Report*. (2013b). Which business school graduates have the most debt? Retrieved from http://premium.usnews.com/best-graduate-schools/top-business-schools/debt-rankings

Usher, A., & Savino, M. (2006). *A world of difference: A global survey of university league tables.* Toronto, ON: Educational Policy Institute. Retrieved from http://www.educationalpolicy.org/pdf/world-of-difference-200602162.pdf

Waldfogel, J., & Chen, L. (2006). Does information undermine brand? Information intermediary use and preference for branded web retailers. *The Journal of Industrial Economics, 54*(4), 425-449. doi:10.1111/j.1467-6451.2006.00295.x

Webb, N. M., Shavelson, R. J., & Haertel, E. H. (2006). Reliability coefficients and generalizability theory. In C. R. Rao & S. Sinharay (Eds.), *Handbook of statistics: Vol. 26. Psychometrics* (pp. 81-124). Amsterdam, The Netherlands: Elsevier B. V. doi:10.1016/S0169-7161(06)26004-8.

Weerts, D. J., & Ronca, J. M. (2006). Examining differences in state support for higher education: A comparative study of state appropriations for research I universities. *The Journal of Higher Education, 77*(6), 935-965.

Weerts, D. J., & Ronca, J. M. (2012). Understanding differences in state support for higher education across states, sectors, and institutions: A longitudinal study. *The Journal of Higher Education, 83*(2), 155-185. doi:10.1353/jhe.2012.0012

Weerts, D. J., Sanford, T., & Reinert, L. (2012). *College funding in context: Understanding the difference in higher education appropriations across the states.* Retrieved from http://www.demos.org/sites/default/files/publications/HigherEducationReportDemos.pdf

Weick, K. E. (1976). Educational organizations as loosely coupled systems. *Administrative Science Quarterly, 21*(1), 1-19. Retrieved from http://www.jstor.org/stable/2391875

Weidlich, T. (2001, April 4). M.B.A. programs that look a lot like work. *New York Times,* p. C9.

Weigley, S., & Hess, A. E. M. (2013, April 25). Universities getting the most government money. [24/7 Wall St. web blog post]. 24/7 Wall St., LLC. Retrieved from http://247wallst.com/special-report/2013/04/25/universities-getting-the-most-government-money

Weiss, R. E. (2010). *Modeling longitudinal data.* New York, NY: Springer Science+Business Media, Inc.

Wharton, University of Pennsylvania. (2009). *MBA for executives: Earn the Wharton MBA in San Francisco.* Retrieved from http://www.wharton.upenn.edu/mbaexecutive/san_francisco/index.html

Whitman, D. (2002, April). Doing the right thing. *The Newsletter of the Association of American Law Schools, President's Message*. Retrieved from http://www.aals.org/presidentsmessages/pmapr02.html

Widaman, K. F. (1985). Hierarchically nested covariance structure models for multitrait-multimethod data. *Applied Psychological Measurement, 9*(1), 1-26. doi:10.1177/014662168500900101

Wiley, M. G., & Zald, M. N. (1968). The growth and transformation of educational accrediting agencies: An exploratory study in social control of institutions.

*Sociology of Education, 41*(1), 36-56. Retrieved from http://www.jstor.org/stable/2112084

Wilson, R. (1985). Reputations in games and markets. In A. E. Roth (Ed.), *Game theoretic models of bargaining* (pp. 65-84). New York, NY: Cambridge University Press.

Zheng, H. Y., & Stewart, A. A. (2000, May 21-23). *Assessing the effectiveness of public research universities: Using NSF/NCES data and data envelopment analysis.* Paper presented at the Annual Forum of the Association for Institutional Research, Cincinnati, OH. Retrieved from http://www.eric.ed.gov/PDFS/ED445642.pdf

Zumel, N., & Mount, J. (2014). *Practical data science with R.* Greenwich, CT: Manning Publications Co.

Zumeta, W. (2001). Public policy and accountability in higher education: Lessons from the past and present for the new millennium. In D. E. Heller (Ed.), *The states and public higher education policy: Affordability, access, and accountability* (pp. 155-197). Baltimore and London: The Johns Hopkins University Press.

Appendix A

Results for Five Major Ranking Systems of M.B.A. Programs in the U.S.A.: Top 30 Programs

Table A1

*2011 U.S. News & World Report. Best Graduate Schools: Schools of Business*

| Rank | School | Overall score | Peer assessment score (5.0= highest) | Recruiter assessment score (5.0= highest) | Full-time average undergrad GPA | Full-time average GMAT score | Full-time acceptance rate | Average starting salary and bonus | Graduates employed at graduation | Employed 3 months later after graduation |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Stanford University (CA) | 100 | 4.8 | 4.6 | 3.69 | 728 | 6.8% | $131,949 | 75.8% | 92.4% |
| 2 | Harvard University (MA) | 98 | 4.8 | 4.5 | 3.67 | 724 | 11.2% | $131,759 | 78.6% | 90.1% |
| 3 | Massachusetts Institute of Technology (Sloan) | 93 | 4.7 | 4.4 | 3.57 | 718 | 13.0% | $125,905 | 75.6% | 88.9% |
| 3 | University of Pennsylvania (Wharton) | 93 | 4.8 | 4.5 | 3.50 | 718 | 16.8% | $132,579 | 72.5% | 84.2% |
| 5 | Northwestern University (Kellogg) (IL) | 92 | 4.7 | 4.5 | 3.52 | 714 | 19.9% | $123,996 | 75.2% | 86.5% |
| 5 | University of Chicago (Booth) | 92 | 4.7 | 4.4 | 3.52 | 715 | 22.3% | $126,779 | 75.8% | 88.6% |
| 7 | Dartmouth College (Tuck) (NH) | 88 | 4.2 | 4.2 | 3.50 | 716 | 20.3% | $128,013 | 80.0% | 93.3% |
| 7 | University of California-Berkeley (Haas) | 88 | 4.5 | 4.1 | 3.63 | 718 | 11.6% | $120,164 | 64.9% | 87.0% |
| 9 | Columbia University (NY) | 87 | 4.4 | 4.3 | 3.50 | 712 | 15.3% | $123,486 | 67.2% | 89.6% |

(continued)

Table A1 (continued)

| Rank | School | Overall score | Peer assessment score (5.0= highest) | Recruiter assessment score (5.0= highest) | Full-time average undergrad GPA | Full-time average GMAT score | Full-time acceptance rate | Average starting salary and bonus | Graduates employed at graduation | Employed 3 months later after graduation |
|---|---|---|---|---|---|---|---|---|---|---|
| 10 | New York University (Stern) | 81 | 4.1 | 4.0 | 3.42 | 715 | 13.2% | $121,867 | 73.7% | 86.4% |
| 10 | Yale University (CT) | 81 | 4.1 | 4.0 | 3.52 | 722 | 17.3% | $113,226 | 64.1% | 89.5% |
| 12 | Duke University (Fuqua) (NC) | 78 | 4.3 | 4.1 | 3.40 | 697 | 23.7% | $118,923 | 63.6% | 76.9% |
| 13 | University of Virginia (Darden) | 77 | 4.0 | 4.1 | 3.41 | 699 | 26.3% | $119,278 | 69.0% | 82.7% |
| 14 | University of California-Los Angeles (Anderson) | 76 | 4.1 | 3.9 | 3.53 | 710 | 29.0% | $108,806 | 54.2% | 80.6% |
| 14 | University of Michigan-Ann Arbor (Ross) | 76 | 4.3 | 4.0 | 3.40 | 704 | 25.4% | $116,201 | 59.8% | 72.1% |
| 16 | Cornell University (Johnson) (NY) | 74 | 4.1 | 4.1 | 3.25 | 687 | 23.4% | $112,039 | 66.8% | 85.9% |
| 17 | University of Texas-Austin (McCombs) | 72 | 3.9 | 3.7 | 3.43 | 684 | 24.0% | $108,886 | 64.8% | 89.7% |
| 18 | Carnegie Mellon University (Tepper) (PA) | 69 | 3.9 | 3.6 | 3.23 | 694 | 27.5% | $106,066 | 69.7% | 88.9% |
| 19 | U. of North Carolina-Chapel Hill (Kenan-Flagler) | 68 | 3.8 | 3.9 | 3.30 | 686 | 36.1% | $111,327 | 60.3% | 78.6% |
| 20 | Washington University in St. Louis (Olin) | 67 | 3.6 | 3.5 | 3.48 | 695 | 28.6% | $90,767 | 61.0% | 92.7% |
| 21 | University of Minnesota-Twin Cities (Carlson | 66 | 3.6 | 3.5 | 3.50 | 694 | 30.3% | $97,298 | 63.0% | 82.7% |

(continued)

Table A1 (continued)

| Rank | School | Overall score | Peer assessment score (5.0= highest) | Recruiter assessment score (5.0= highest) | Full-time average undergrad GPA | Full-time average GMAT score | Full-time acceptance rate | Average starting salary and bonus | Graduates employed at graduation | Employed 3 months later after graduation |
|---|---|---|---|---|---|---|---|---|---|---|
| 23 | Emory University (Goizueta) (GA) | 65 | 3.7 | 3.5 | 3.34 | 680 | 33.3% | $100,300 | 57.2% | 87.3% |
| 23 | Indiana University-Bloomington (Kelley) | 65 | 3.7 | 3.7 | 3.31 | 664 | 41.7% | $101,206 | 70.1% | 83.8% |
| 25 | Georgetown University (McDonough) (DC) | 64 | 3.5 | 3.4 | 3.37 | 684 | 42.5% | $103,676 | 58.0% | 88.1% |
| 25 | Ohio State University (Fisher) | 64 | 3.6 | 3.3 | 3.40 | 676 | 27.0% | $91,628 | 62.5% | 94.6% |
| 27 | Arizona State University (Carey) | 61 | 3.5 | 3.4 | 3.31 | 672 | 29.6% | $92,101 | 62.2% | 89.2% |
| 28 | Georgia Institute of Technology | 60 | 3.3 | 3.2 | 3.28 | 678 | 22.0% | $92,282 | 67.7% | 95.2% |
| 28 | University of California-Davis | 60 | 3.1 | 3.4 | 3.27 | 692 | 23.9% | $96,295 | 67.4% | 93.0% |
| 28 | University of Wisconsin-Madison | 60 | 3.5 | 3.3 | 3.36 | 675 | 30.5% | $93,332 | 55.0% | 87.4% |
| 28 | Vanderbilt University (Owen) (TN) | 60 | 3.4 | 3.4 | 3.40 | 673 | 36.0% | $93,351 | 62.9% | 82.9% |
| 32 | Brigham Young University (Marriott) (UT) | 58 | 3.0 | 3.2 | 3.50 | 675 | 44.3% | $97,207 | 62.2% | 86.6% |

*Source. U.S. News & World Report*. (2011). *U.S. News & World Report best graduate schools guidebook* (2012 edition). Washington, D.C.: *U.S. News & World Report* L.P.

Table A2

*Bloomberg Businessweek. The Best U.S. B-Schools of 2010*

| 2010 Rank | School | Graduate poll | Corporate poll | Intellectual capital | Index number |
|---|---|---|---|---|---|
| 1 | Chicago (Booth) | 2 | 1 | 8 | 100 |
| 2 | Harvard | 5 | 3 | 14 | 95.6 |
| 3 | Pennsylvania (Wharton) | 11 | 2 | 13 | 94.4 |
| 4 | Northwestern (Kellogg) | 6 | 4 | 28 | 92.5 |
| 5 | Stanford | 7 | 7 | 7 | 87.4 |
| 6 | Duke (Fuqua) | 15 | 8 | 1 | 86.9 |
| 7 | Michigan (Ross) | 16 | 5 | 11 | 85.5 |
| 8 | UC-Berkeley (Haas) | 4 | 11 | 5 | 84.9 |
| 9 | Columbia | 10 | 9 | 12 | 84.5 |
| 10 | MIT (Sloan) | 3 | 14 | 18 | 81.3 |
| 11 | Virginia (Darden) | 1 | 16 | 36 | 81.0 |
| 12 | Southern Methodist (Cox) | 21 | 6 | 40 | 79.5 |
| 13 | Cornell (Johnson) | 8 | 23 | 6 | 77.1 |
| 14 | Dartmouth (Tuck) | 14 | 15 | 4 | 77.0 |
| 15 | Carnegie Mellon (Tepper) | 12 | 20 | 16 | 75.2 |
| 16 | North Carolina (Kenan-Flagler) | 19 | 17 | 26 | 72.4 |
| 17 | UCLA (Anderson) | 17 | 26 | 15 | 71.4 |
| 18 | NYU (Stern) | 20 | 22 | 19 | 70.5 |

(continued)

Table A2 (continued)

| 2010 Rank | School | Graduate poll | Corporate poll | Intellectual capital | Index number |
|---|---|---|---|---|---|
| 19 | Indiana (Kelley) | 26 | 18 | 46 | 67.2 |
| 20 | Michigan State (Broad) | 13 | 32 | 39 | 66.8 |
| 21 | Yale | 18 | 37 | 9 | 66.1 |
| 22 | Emory (Goizueta) | 22 | 31 | 21 | 65.3 |
| 23 | Georgia Tech | 25 | 30 | 20 | 64.3 |
| 24 | Notre Dame (Mendoza) | 24 | 33 | 25 | 63.8 |
| 25 | Texas-Austin (McCombs) | 29 | 29 | 17 | 63.5 |
| 26 | USC (Marshall) | 9 | 45 | 35 | 61.6 |
| 27 | Brigham Young (Marriott) | 31 | 19 | 49 | 60.8 |
| 28 | Minnesota (Carlson) | 35 | 21 | 29 | 58.3 |
| 29 | Rice (Jones) | 27 | 44 | 10 | 58.0 |
| 30 | Texas A&M (Mays) | 43 | 13 | 34 | 57.8 |

*Source*. *Bloomberg Businessweek.* (2011). *The best U.S. B-schools of 2010.* Retrieved from http://www.businessweek.com/interactive_reports/bs_2010_US_FTMBA_TAB_1111.html

Table A3

*The Financial Times Global MBA Rankings 2011*

| Rank 2011 | School Name | Country | Weighted salary (US$) | Salary percentage increase | Value for money rank | Career progress rank | Aims achieved rank | Placement success rank | Employed at three months (%) | Alumni recommend rank | Women faculty (%) | Women students (%) | Women board (%) | International faculty (%) | International students (%) | International board (%) | International mobility rank | International experience rank | Languages | Faculty with doctorates (%) | FT doctoral rank | FT research rank |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | London Business School | U.K | 145,776 | 132 | 57 | 11 | 2 | 41 | 91 | 4 | 24 | 28 | 16 | 85 | 92 | 60 | 2 | 11 | 1 | 98 | 23 | 7 |
| 1 | University of Pennsylvania: Wharton | U.S.A. | 171,551 | 123 | 89 | 33 | 32 | 11 | 84 | 2 | 19 | 40 | 8 | 37 | 34 | 66 | 43 | 33 | 0 | 100 | 2 | 2 |
| 3 | Harvard Business School | U.S.A. | 170,238 | 116 | 87 | 22 | 51 | 32 | 90 | 1 | 22 | 36 | 48 | 37 | 34 | 21 | 52 | 59 | 0 | 92 | 14 | 1 |
| 4 | Insead | France / Singapore | 147,883 | 108 | 4 | 18 | 14 | 44 | 82 | 7 | 14 | 33 | 17 | 90 | 92 | 80 | 8 | 5 | 2 | 97 | 17 | 10 |
| 4 | Stanford Graduate School of Business | U.S.A. | 183,260 | 115 | 98 | 9 | 17 | 20 | 92 | 3 | 19 | 39 | 17 | 38 | 41 | 25 | 56 | 15 | 0 | 92 | 4 | 8 |
| 6 | Hong Kong UST Business School | China | 133,334 | 142 | 16 | 36 | 31 | 52 | 69 | 50 | 25 | 34 | 39 | 88 | 93 | 94 | 19 | 9 | 1 | 100 | 19 | 14 |

(continued)

Table A3 (continued)

| Rank 2011 | School Name | Country | Weighted salary (US$) | Salary percentage increase | Value for money rank | Career progress rank | Aims achieved rank | Placement success rank | Employed at three months (%) | Alumni recommend rank | Women faculty (%) | Women students (%) | Women board (%) | International faculty (%) | International students (%) | International board (%) | International mobility rank | International experience rank | Languages | Faculty with doctorates (%) | FT doctoral rank | FT research rank |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 7 | Columbia Business School | U.S.A. | 163,407 | 117 | 97 | 17 | 28 | 23 | 90 | 6 | 18 | 35 | 10 | 62 | 46 | 35 | 49 | 57 | 0 | 97 | 7 | 10 |
| 8 | IE Business School | Spain | 149,584 | 136 | 28 | 4 | 82 | 49 | 91 | 29 | 33 | 29 | 23 | 54 | 87 | 82 | 31 | 34 | 1 | 93 | 69 | 62 |
| 9 | Iese Business School | Spain | 131,890 | 138 | 65 | 3 | 8 | 25 | 94 | 17 | 18 | 24 | 28 | 49 | 80 | 90 | 3 | 4 | 1 | 100 | 48 | 70 |
| 9 | MIT: Sloan | U.S.A. | 158,387 | 121 | 88 | 10 | 54 | 16 | 89 | 8 | 22 | 35 | 13 | 32 | 51 | 60 | 51 | 25 | 0 | 100 | 11 | 19 |
| 11 | Indian Institute of Management, Ahmedabad | India | 174,440 | 152 | 26 | 1 | 94 | 14 | 97 | 10 | 12 | 7 | 30 | 0 | 8 | 0 | 42 | 81 | 0 | 96 | 55 | 92 |
| 12 | University of Chicago: Booth | U.S.A. | 151,373 | 109 | 95 | 55 | 26 | 1 | 89 | 9 | 15 | 35 | 13 | 37 | 44 | 38 | 65 | 50 | 0 | 96 | 8 | 4 |
| 13 | Indian School of Business | India | 134,406 | 187 | 39 | 28 | 66 | 35 | 98 | 28 | 15 | 28 | 5 | 21 | 5 | 55 | 35 | 77 | 0 | 100 | n/a | 81 |
| 14 | IMD | Switzerland | 145,846 | 89 | 6 | 5 | 1 | 2 | 83 | 15 | 10 | 23 | 16 | 98 | 100 | 81 | 1 | 46 | 1 | 100 | n/a | 67 |
| 15 | New York University: Stern | U.S.A. | 138,865 | 119 | 91 | 34 | 40 | 10 | 86 | 11 | 16 | 36 | 10 | 49 | 35 | 16 | 59 | 31 | 0 | 100 | 5 | 13 |

(continued)

Table A3 (continued)

| Rank 2011 | School Name | Country | Weighted salary (US$) | Salary percentage increase | Value for money rank | Career progress rank | Aims achieved rank | Placement success rank | Employed at three months (%) | Alumni recommend rank | Women faculty (%) | Women students (%) | Women board (%) | International faculty (%) | International students (%) | International board (%) | International mobility rank | International experience rank | Languages | Faculty with doctorates (%) | FT doctoral rank | FT research rank |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 15 | Yale School of Management | U.S.A. | 146,959 | 133 | 79 | 41 | 6 | 30 | 90 | 20 | 19 | 37 | 24 | 31 | 33 | 17 | 72 | 35 | 0 | 97 | 35 | 33 |
| 17 | Ceibs | China | 126,315 | 155 | 46 | 43 | 48 | 69 | 92 | 40 | 13 | 37 | 17 | 65 | 42 | 50 | 50 | 22 | 1 | 96 | n/a | 76 |
| 18 | Dartmouth College: Tuck | U.S.A. | 155,020 | 113 | 90 | 56 | 3 | 15 | 93 | 13 | 21 | 35 | 17 | 31 | 40 | 12 | 55 | 38 | 0 | 100 | n/a | 22 |
| 18 | HEC Paris | France | 122,828 | 106 | 36 | 14 | 4 | 66 | 89 | 27 | 27 | 26 | 7 | 65 | 83 | 41 | 5 | 13 | 1 | 99 | 47 | 33 |
| 20 | Duke University: Fuqua | U.S.A. | 136,563 | 107 | 93 | 68 | 22 | 9 | 77 | 16 | 18 | 30 | 10 | 39 | 39 | 41 | 74 | 28 | 0 | 99 | 20 | 3 |
| 21 | Esade Business School | Spain | 125,346 | 128 | 40 | 2 | 41 | 45 | 82 | 35 | 29 | 31 | 25 | 19 | 85 | 90 | 12 | 17 | 1 | 82 | 66 | 89 |
| 21 | Northwestern University: Kellogg | U.S.A. | 143,365 | 100 | 100 | 27 | 19 | 7 | 88 | 5 | 19 | 31 | 8 | 45 | 34 | 10 | 70 | 16 | 0 | 95 | 6 | 8 |
| 23 | National University of Singapore School of Business | Singapore | 100,456 | 140 | 35 | 53 | 62 | 64 | 93 | 56 | 31 | 33 | 17 | 59 | 88 | 17 | 9 | 12 | 0 | 97 | 51 | 43 |

(continued)

Table A3 (continued)

| Rank 2011 | School Name | Country | Weighted salary (US$) | Salary percentage increase | Value for money rank | Career progress rank | Aims achieved rank | Placement success rank | Employed at three months (%) | Alumni recommend rank | Women faculty (%) | Women students (%) | Women board (%) | International faculty (%) | International students (%) | International board (%) | International mobility rank | International experience rank | Languages | Faculty with doctorates (%) | FT doctoral rank | FT research rank |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 24 | University of Michigan: Ross | U.S.A. | 137,189 | 104 | 78 | 37 | 39 | 4 | 78 | 14 | 23 | 30 | 20 | 31 | 29 | 17 | 58 | 44 | 0 | 94 | 3 | 6 |
| 25 | University of California at Berkeley: Haas | U.S.A. | 144,790 | 87 | 85 | 47 | 23 | 6 | 87 | 12 | 30 | 30 | 9 | 44 | 42 | 9 | 61 | 30 | 0 | 100 | 9 | 5 |
| 26 | University of Cambridge: Judge | U.K. | 137,199 | 101 | 18 | 20 | 18 | 55 | 74 | 43 | 10 | 26 | 19 | 52 | 94 | 55 | 11 | 70 | 0 | 96 | 52 | 51 |
| 27 | University of Oxford: Saïd | U.K. | 132,905 | 102 | 19 | 30 | 42 | 73 | 71 | 30 | 15 | 24 | 30 | 56 | 95 | 32 | 4 | 56 | 0 | 97 | 32 | 51 |
| 28 | SDA Bocconi | Italy | 110,186 | 123 | 12 | 32 | 71 | 59 | 88 | 51 | 40 | 33 | 21 | 29 | 67 | 64 | 22 | 43 | 0 | 88 | 12 | 62 |
| 29 | Manchester Business School | U.K. | 116,100 | 111 | 48 | 13 | 27 | 70 | 80 | 39 | 33 | 23 | 23 | 35 | 93 | 23 | 14 | 19 | 0 | 89 | 1 | 70 |
| 30 | Cornell University: Johnson | U.S.A. | 140,273 | 107 | 92 | 46 | 16 | 13 | 81 | 25 | 25 | 29 | 16 | 33 | 31 | 27 | 64 | 42 | 0 | 89 | 65 | 33 |

*Source. The Financial Times (2011). Global MBA Rankings 2011.* Retrieved from http://rankings.ft.com/exportranking/global-mba-rankings-2011/pdf

Table A4

*The Economist. 2011 Full-Time MBA Ranking: United States*

| Overall Rank | Regional Rank | Business School | Open new career opportunities rank (out of 100) | Personal development and educational experience rank (out of 100) | Increase in salary rank (out of 100) | Potential to network rank (out of 100) | Student rating of program (out of 5) |
|---|---|---|---|---|---|---|---|
| 1 | 1 | Dartmouth College - Tuck School of Business | 2 | 1 | 19 | 28 | 4.6 |
| 2 | 2 | Chicago, University of - Booth School of Business | 1 | 7 | 32 | 20 | 4.7 |
| 4 | 3 | Virginia, University of - Darden Graduate School of Business Administration | 4 | 3 | 34 | 11 | 4.8 |
| 5 | 4 | Harvard Business School | 9 | 20 | 15 | 14 | 4.3 |
| 6 | 5 | California at Berkeley, University of - Haas School of Business | 8 | 14 | 23 | 9 | 4.2 |
| 7 | 6 | Columbia Business School | 3 | 13 | 31 | 24 | 4.4 |
| 8 | 7 | Stanford Graduate School of Business | 25 | 2 | 10 | 16 | 4.6 |
| 11 | 8 | Massachusetts Institute of Technology - MIT Sloan School of Management | 11 | 5 | 22 | 39 | 4.5 |
| 12 | 9 | New York University - Leonard N Stern School of Business | 14 | 10 | 32 | 6 | 4.6 |
| 15 | 10 | Pennsylvania, University of - Wharton School | 22 | 16 | 18 | 22 | 4.1 |
| 16 | 11 | Carnegie Mellon University - The Tepper School of Business | 10 | 17 | 33 | 42 | 4.7 |
| 18 | 12 | Northwestern University - Kellogg School of Management | 21 | 18 | 25 | 12 | 4.4 |
| 20 | 13 | Duke University - Fuqua School of Business | 34 | 11 | 26 | 48 | 4.3 |
| 22 | 14 | Southern California, University of - Marshall School of Business | 6 | 64 | 68 | 8 | 4.2 |
| 25 | 15 | Cornell University – Samuel Curtis Johnson Graduate School of Management | 23 | 24 | 42 | 34 | 4.5 |
| 26 | 16 | Yale School of Management | 37 | 9 | 36 | 64 | 4.5 |

(continued)

Table A4 (continued)

| Overall Rank | Regional Rank | Business School | Open new career opportunities rank (out of 100) | Personal development and educational experience rank (out of 100) | Increase in salary rank (out of 100) | Potential to network rank (out of 100) | Student rating of program (out of 5) |
|---|---|---|---|---|---|---|---|
| 27 | 17 | UCLA Anderson School of Management | 29 | 22 | 48 | 40 | 4.5 |
| 28 | 18 | Emory University - Goizueta Business School | 26 | 29 | 41 | 37 | 4.4 |
| 29 | 19 | Hult International Business School | 31 | 75 | 7 | 63 | 4.3 |
| 30 | 20 | Michigan, University of - Stephen M. Ross School of Business | 43 | 28 | 38 | 27 | 4.5 |
| 31 | 21 | Texas at Austin, University of - McCombs School of Business | 16 | 45 | 57 | 36 | 4.3 |
| 34 | 22 | Vanderbilt University - Owen Graduate School of Management | 20 | 25 | 74 | 59 | 4.6 |
| 35 | 23 | Washington, University of--Foster School of Business | 18 | 32 | 78 | 35 | 4.2 |
| 39 | 24 | Indiana University - Kelley School of Business | 33 | 58 | 56 | 32 | 4.5 |
| 40 | 25 | Notre Dame, University of - Mendoza College of Business | 45 | 73 | 50 | 10 | 4.4 |
| 42 | 26 | Washington University in St Louis - Olin Business School | 42 | 66 | 77 | 56 | 4.5 |
| 44 | 27 | Georgetown University - Robert Emmett McDonough School of Business | 36 | 54 | 61 | 45 | 4.2 |
| 45 | 28 | Wisconsin School of Business | 17 | 65 | 81 | 50 | 4.4 |
| 47 | 28 | North Carolina at Chapel Hill, University of - Kenan-Flagler Business School | 56 | 34 | 51 | 53 | 4.4 |
| 48 | 30 | Wake Forest University Schools of Business | 30 | 69 | 72 | 38 | 4.3 |

*Source. The Economist* (2012a, March 13). *2011 Full-Time MBA Ranking: United States.* Retrieved from http://www.economist.com/whichmba/full-time-mba-ranking?term_node_tid_depth=77633

Table A5

*Forbes. The Best Business Schools 2011*

| Rank | School | Total 5-year MBA gain ($thou) Class of 2006 5-year total compensation after graduation, minus the sum of tuition, fees and forgone compensation | 5-year MBA gain as % of expenses | Years to payback | Pre-MBA salary ($thou) | 2010 Salary ($thou) | Tuition ($thou) |
|---|---|---|---|---|---|---|---|
| 1 | Harvard | 118 | 51 | 3.6 | 79 | 230 | 112 |
| 2 | Stanford | 116 | 54 | 3.5 | 77 | 205 | 110 |
| 3 | Chicago (Booth) | 112 | 52 | 3.5 | 72 | 205 | 106 |
| 4 | Pennsylvania (Wharton) | 111 | 48 | 3.7 | 79 | 225 | 109 |
| 5 | Columbia | 106 | 48 | 3.6 | 76 | 198 | 115 |
| 6 | Dartmouth (Tuck) | 102 | 48 | 3.6 | 74 | 200 | 104 |
| 7 | Northwestern (Kellogg) | 101 | 50 | 3.5 | 68 | 183 | 105 |
| 8 | Cornell (Johnson) | 100 | 53 | 3.5 | 65 | 180 | 104 |
| 9 | Virginia (Darden) | 99 | 53 | 3.5 | 68 | 178 | 104 |
| 10 | MIT (Sloan) | 97 | 42 | 3.7 | 77 | 200 | 104 |
| 11 | Yale | 96 | 54 | 3.3 | 57 | 150 | 105 |
| 12 | Duke (Fuqua) | 93 | 49 | 3.6 | 63 | 162 | 102 |
| 13 | UC Berkeley (Haas) | 88 | 49 | 3.7 | 68 | 171 | 102 |
| 14 | Michigan (Ross) | 81 | 45 | 3.6 | 62 | 146 | 103 |
| 15 | Brigham Young (Marriott) | 81 | 73 | 3.1 | 47 | 108 | 42 |
| 16 | UNC (Kenan-Flagler) | 78 | 48 | 3.6 | 55 | 136 | 96 |
| 17 | Texas-Austin (McCombs) | 78 | 47 | 3.6 | 59 | 144 | 93 |
| 18 | NYU (Stern) | 77 | 42 | 3.7 | 61 | 160 | 97 |
| 19 | Michigan State (Broad) | 77 | 61 | 3.1 | 45 | 109 | 77 |
| 20 | UCLA (Anderson) | 77 | 45 | 3.7 | 63 | 159 | 100 |
| 21 | Iowa (Tippie) | 76 | 75 | 3.1 | 36 | 96 | 68 |
| 22 | Emory (Goizueta) | 72 | 47 | 3.5 | 55 | 128 | 86 |
| 23 | Carnegie Mellon (Tepper) | 72 | 38 | 3.8 | 55 | 147 | 105 |

(continued)

Table A5 (continued)

| Rank | School | Total 5-year MBA gain ($thou) Class of 2006 5-year total compensation after graduation, minus the sum of tuition, fees and forgone compensation | 5-year MBA gain as % of expenses | Years to payback | Pre-MBA salary ($thou) | 2010 Salary ($thou) | Tuition ($thou) |
|---|---|---|---|---|---|---|---|
| 24 | Texas A&M (Mays) | 72 | 64 | 3.3 | 46 | 116 | 43 |
| 25 | SMU (Cox) | 71 | 53 | 3.4 | 45 | 112 | 90 |
| 26 | Penn State (Smeal) | 70 | 56 | 3.2 | 49 | 111 | 71 |
| 27 | Indiana (Kelley) | 69 | 51 | 3.3 | 47 | 109 | 90 |
| 28 | Minnesota (Carlson) | 69 | 46 | 3.4 | 52 | 119 | 89 |
| 29 | Wisconsin-Madison | 69 | 53 | 3.5 | 50 | 119 | 53 |
| 30 | Wake Forest (Babcock) | 66 | 50 | 3.4 | 45 | 107 | 77 |

*Source. Forbes.* (2011). *The best business schools.* Retrieved from http://www.forbes.com/lists/2011/95/best-business-schools-11_rank.htm

# Appendix B

## Berlin Principles on Ranking of Higher Education Institutions

Rankings and league tables should:

### A) Purpose and Goals of Rankings

1. ***Be one of a number of diverse approaches to the assessment of higher education inputs, processes, and outputs.*** Rankings can provide comparative information and improved understanding of higher education, but should not be the main method for assessing what higher education is and does. Rankings provide a market-based perspective that can complement the work of government, accrediting authorities, and independent review agencies.

2. ***Be clear about their purposes and their target groups.*** Rankings have to be designed with due regard to their purpose. Indicators designed to meet a particular objective or to inform one target group may not be adequate for different purposes or target groups.

3. ***Recognize the diversity of institutions and take the different missions and goals of institutions into account.*** Quality measures for research-oriented institutions, for example, are quite different from those that are appropriate for institutions that provide broad access to underserved communities. Institutions that are being ranked and the experts that inform the ranking process should be consulted often.

4. ***Provide clarity about the range of information sources for rankings and the messages each source generates.*** The relevance of ranking results depends on the audiences receiving the information and the sources of that information (such as databases, students, professors, employers). Good practice would be to combine the different perspectives provided by those sources in order to get a more complete view of each higher education institution included in the ranking.

5. ***Specify the linguistic, cultural, economic, and historical contexts of the educational system being ranked.*** International rankings in particular should be aware of possible biases and be precise about their objective. Not all nations or systems share the same values and beliefs about what constitutes "quality" in tertiary institutions, and ranking systems should not be devised to force such comparisons.

### B) Design and Weighting of Indicators

6. ***Be transparent regarding the methodology used for creating the rankings.*** The choice of methods used to prepare rankings should be clear and unambiguous. This transparency should include the calculation of indicators as well as the origin of data.

7. ***Choose indicators according to their relevance and validity.*** The choice of data should be grounded in recognition of the ability of each measure to represent quality and academic and institutional strengths, and not availability of data. Be clear about why measures were included and what they are meant to represent.

8. ***Measure outcomes in preference to inputs whenever possible.*** Data on inputs are relevant as they reflect the general condition of a given establishment and are more frequently available. Measures of outcome provide a more accurate assessment of the standing and/or quality of a given institution or program, and compilers of rankings should ensure that an appropriate balance is achieved.

9. ***Make the weights assigned to different indicators (if used) prominent and limit changes to them.*** Changes in weights make it difficult for consumers to discern whether an institution's or program's status changed in the rankings due to an inherent difference or due to a methodological change.

**C) Collection and Processing of Data**

10. ***Pay due attention to ethical standards and the good practice recommendations articulated in these Principles.*** In order to assure the credibility of each ranking, those responsible for collecting and using data and undertaking on-site visits should be as objective and impartial as possible.

11. ***Use audited and verifiable data whenever possible.*** Such data have several advantages, including the fact that they have been accepted by institutions and that they are comparable and compatible across institutions.

12. ***Include data that are collected with proper procedures for scientific data collection.*** Data collected from an unrepresentative or skewed subset of students, faculty, or other parties may not accurately represent an institution or program and should be excluded.

13. ***Apply measures of quality assurance to ranking processes themselves.*** These processes should take note of the expertise that is being applied to evaluate institutions and use this knowledge to evaluate the ranking itself. Rankings should be learning systems continuously utilizing this expertise to develop methodology.

14. ***Apply organizational measures that enhance the credibility of rankings.*** These measures could include advisory or even supervisory bodies, preferably with some international participation.

**D) Presentation of Ranking Results**

15. ***Provide consumers with a clear understanding of all of the factors used to develop a ranking, and offer them a choice in how rankings are displayed.*** This way, the users of rankings would have a better understanding of the indicators that are used to rank institutions or programs. In addition, they should have some opportunity to make their own decisions about how these indicators should be weighted.

16. ***Be compiled in a way that eliminates or reduces errors in original data, and be organized and published in a way that errors and faults can be corrected.*** Institutions and the public should be informed about errors that have occurred.

Berlin, 20 May 2006

*Source.* The International Ranking Expert Group. (2006). *Berlin Principles on Ranking of Higher Education Institutions.* Berlin, Germany: The International Ranking Expert Group. Retrieved from http://www.che.de/doenloads/Berlin_Principles_IREG-534.pdf